

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 22:47:08 ; Search time 3292.45 Seconds  
(without alignments)  
3979.632 Million cell updates/sec

Title: US-09-428-122-1

Perfect score: 3000  
Sequence: 1 cgcctcctacctaagtaggtc.....tgatttgctaaacaccttc 3000

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1033670 seqs, 2183789903 residues

Total number of hits satisfying chosen parameters: 2067340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba1: \*  
2: gb\_ba2: \*  
3: gb\_om: \*  
4: gb\_ov: \*  
5: gb\_pat: \*  
6: gb\_ph: \*  
7: gb\_pl1: \*  
8: gb\_pl2: \*  
9: gb\_pr1: \*  
10: gb\_pr2: \*  
11: gb\_pr3: \*  
12: em\_fun: \*  
13: em\_hum1: \*  
14: em\_hum2: \*  
15: em\_in: \*  
16: em\_om: \*  
17: em\_or: \*  
18: em\_ov: \*  
19: em\_pat: \*  
20: em\_ph: \*  
21: em\_pl: \*  
22: em\_ro: \*  
23: em\_sts: \*  
24: em\_sy: \*  
25: em\_un: \*  
26: em\_vl: \*  
27: gb\_hcg1: \*  
28: gb\_hcg2: \*  
29: gb\_in1: \*  
30: gb\_in2: \*  
31: em\_ba1: \*  
32: em\_ba2: \*  
33: em\_hum3: \*  
34: em\_hum4: \*  
35: gb\_pr4: \*  
36: gb\_hcg3: \*  
37: gb\_hcg4: \*  
38: gb\_hcg5: \*  
39: gb\_hcg6: \*  
40: gb\_hcg7: \*  
41: em\_hcg1: \*  
42: em\_hcg2: \*  
43: em\_hcg3: \*

44: em\_hum5: \*  
45: gb\_pl3: \*  
46: gb\_pr5: \*  
47: gb\_hcg8: \*  
48: gb\_hcg9: \*  
49: gb\_hcg10: \*  
50: gb\_hcg11: \*  
51: gb\_hcg12: \*  
52: gb\_hcg13: \*  
53: gb\_hcg14: \*  
54: gb\_in3: \*  
55: gb\_hcg15: \*  
56: gb\_hcg16: \*  
57: gb\_hcg17: \*  
58: em\_hcg4: \*  
59: em\_hcg5: \*  
60: em\_hcg6: \*  
61: em\_hcg7: \*  
62: em\_hum6: \*  
63: gb\_hcg18: \*  
64: gb\_hcg19: \*  
65: gb\_ba3: \*  
66: em\_hcg8: \*  
67: em\_hcg9: \*  
68: em\_hcg10: \*  
69: em\_hcg11: \*  
70: em\_hcg12: \*  
71: em\_hcg13: \*  
72: em\_hcg14: \*  
73: em\_hcg15: \*  
74: em\_hcg16: \*  
75: em\_hcg17: \*  
76: em\_hcg18: \*  
77: em\_hcg19: \*  
78: em\_hcg20: \*  
79: em\_hcg21: \*  
80: em\_hcg22: \*  
81: em\_hcg23: \*  
82: gb\_pr6: \*  
83: gb\_pr7: \*  
84: gb\_hcg20: \*  
85: gb\_hcg21: \*  
86: gb\_hcg22: \*  
87: gb\_hcg23: \*  
88: gb\_ro: \*  
89: gb\_sts1: \*  
90: gb\_sts2: \*  
91: gb\_sy: \*  
92: gb\_un: \*  
93: gb\_vl1: \*  
94: gb\_vl2: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2988	99.6	16448	1	AE001587	AE001587 Chlamydia
2	2988	99.6	26920	2	CPN133035	AJ133035 Chlamydia
3	2986.4	99.5	12127	1	AE002235	AE002235 Chlamydia
4	2986.4	99.5	300050	1	AP002545	AP002545 Chlamydia
5	2775.8	92.5	2787	5	AB1835	AB1835 Sequence 9
6	2239.4	74.6	2241	1	AB033795	AB033795 Chlamydia
7	546.6	18.2	4926	2	CP072499	U72499 Chlamydia
8	531.8	17.7	12676	1	AE002192	AE002192 Chlamydia
9	531.8	17.7	300600	1	AP002546	AP002546 Chlamydia
10	530.2	17.7	10757	1	AE001628	AE001628 Chlamydia
11	530.2	17.7	2815	5	AB1829	AB1829 Sequence 3
12	530.2	17.7	17280	2	CPN133034	AJ133034 Chlamydia

13	523	17.4	534	1	AB033794	AB033794	Chlamydo
14	522.2	17.4	2757	5	AB1837	AB1837	Sequence 11
15	518.4	17.3	10026	1	AE002193	AE002193	Chlamydo
16	518.4	17.3	306500	1	AP002546	AP002546	Chlamydo
17	517.2	17.2	10757	1	AB033802	AB033802	Chlamydo
18	516.6	17.2	10757	1	AE001628	AE001628	Chlamydo
19	515.6	17.2	12800	2	CPN133034	CPN133034	Chlamydo
20	515.2	17.2	15068	1	AE001627	AE001627	Chlamydo
21	515	17.2	2787	1	AB033803	AB033803	Chlamydo
22	515	17.2	2787	5	AB1839	AB1839	Sequence 13
23	515	17.2	12676	1	AE002192	AE002192	Chlamydo
24	514	17.1	2793	5	AB1841	AB1841	Sequence 15
25	506.6	16.9	8658	1	AB038349	AB038349	Chlamydo
26	506	16.9	3000	2	AB1851	AB1851	Sequence 25
27	505	16.9	6030	2	CPOMP54	CPOMP54	Chlamydo
28	465.8	15.5	2787	1	AB033806	AB033806	Chlamydo
29	465.8	15.5	3200	1	AB1827	AB1827	Sequence 1
30	465.8	15.5	6030	2	CPOMP54	CPOMP54	Chlamydo
31	318.4	12.6	2838	5	AB1849	AB1849	Sequence 23
32	349.4	11.6	6110	2	CP065942	CP065942	Chlamydo
33	349.4	11.6	6234	2	CP065943	CP065943	Chlamydo
34	348.8	11.6	10574	1	AE001586	AE001586	Chlamydo
35	342.4	11.4	2525	1	AB033793	AB033793	Chlamydo
36	340.2	11.3	2526	5	AB1833	AB1833	Sequence 7
37	338	11.3	1377	1	AB033804	AB033804	Chlamydo
38	289	9.6	1065	1	AB033797	AB033797	Chlamydo
39	274.2	9.1	10044	1	AE002237	AE002237	Chlamydo
40	273.4	9.1	10766	1	AE001585	AE001585	Chlamydo
41	273.2	9.1	3052	5	AB1831	AB1831	Sequence 5
42	272.2	9.1	2769	1	AB033792	AB033792	Chlamydo
43	243.4	8.1	10695	1	AE001629	AE001629	Chlamydo
44	243.4	8.1	15789	1	AE002191	AE002191	Chlamydo
45	243.2	8.1	2922	1	AB033808	AB033808	Chlamydo

## ALIGNMENTS

## RESULT 1

LOCUS AE001587 16448 bp DNA BCT 08-MAR-1999  
 DEFINITION Chlamydia pneumoniae section 3 of 103 of the complete genome.  
 ACCESSION AE001587 AE001363  
 VERSION AE001587.1 GI:4376271

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

## TITLE

## JOURNAL

## REVIEW

## AUTHORS

QY	421	tgtaaacaagcagcgctggtagataataatctaccgaagttatagggtttcttcgcacatt	480
Db	492	TGTTAAAGCAGCGCTGGTAGTAATATACACAGTTTATAGGGTTTCTTCGGTACTTT	551
QY	481	tattgcctccctcggaaagttccgaataactaccggaagaagccgttagcgcctcagg	540
Db	552	TATTGCGTCTCTGGAAAGTTCCGATTAATACCGGCAAAAGGACCGTTTAGCTGCTACGGG	611
QY	541	taagcttagtttgacaaaaaatgfcagttctgcttcacgcaaaaactttcaacggtaa	600
Db	612	TAGCTTAGTGTGAC-AAAATGTCATTTGCTCTTACGAAAACTTTCAACGGATAA	670
QY	601	tggcggctgctataccgcgcaaaaactcttcaataacagggaactcaaatgfcagctcgt	660
Db	671	TGCGGGTGCATATCACCCCAAAACTTTTCATTAAACGAGGACTAATGTCAGCTGTG	730
QY	661	tttcgaaatatcaccctcccaagaagaagcggagccattcgaacttcgatagccttaacat	720
Db	731	TTTCGAAAAATACCTCCCTCAAGAAAGGCGGAGGCATTCACACTTCGATCCCTTACAT	790
QY	721	tactcgaaacccaaggggaagctctctttcttgacaactcttcggatctcggatctgc	780
Db	791	TACTGGAAACCAAGGGAAGTCTTTTCTTGACATACTCTTCGATTTCTGAGCTGC	850
QY	781	aattttacagaagcctcggctgacatttcttaataatgctcaaaagtcttccttaatgcaa	840
Db	851	AATTTTACAGAGCCCTCGGTGACTATTCTTAATAATGCTAAAGTTTCCTTATTGACA	910
QY	841	taaggttcaacggagcggagctccctcaacaacgggggataatgcaaggatgctatctgtgc	900
Db	911	TAAAGTACAGGAGCGAGCCTCTCAACAAAGGGGATATCTCAGGAGGTGCTATCTGTGC	970
QY	901	ttataaaccagtaacagataactaagatgcaccctcccaacgggaactcagatgtactactcag	960
Db	971	TTATAAACTAGTACATATCTTAAGTGCACCTCTACCTAGGAAATCAGATGTACTCTTCAG	1030
QY	961	caaatatatactgacaaacagcgggagagctatctatgttanaaaagctcgaactgcttc	1020
Db	1031	CAAAATAATACCTGACACACAGCGGAGAGTATCTATGTAATAAAGTGTCAATGTGCTTC	1080

QY	2651	1TTTGCCCTTACCTATCGGATCCGATTGTGANTAGGAATCAGACCTCCGAAGATGCAAG
Db	2651	1TTTGCCCTTACCTATCGGATCCGATTGTGANTAGGAATCAGACCTCCGAAGATGCAAG
QY	2641	caatccactcttggtatctactgtagctctgttcgttagtaaccgactgtagac
Db	2711	CAATCTACCTCTTGTTATACGTGGATCTTGTGATGTAACCCGACTGTGACAC
QY	2701	actcggaattaggggtagtctcttggaacaccttcggtacgaatttggcaacaagcttt
Db	2771	ACTCGGAATTAGCGGCGATCTTGGAAAACTTGGGACGATTTGGCAGACAACTTT
QY	2761	agctctcgtagcagggaaccatttctgcttaactcaaatittgaagcccttgccaatt
Db	2831	AGCTCTCGTAGCAGGGAACCATTTTGGTTTAACTCAATTTTGAAGACCTTTAGCCAA
QY	2821	ttctttgaattggcgtgggtatctctcgcaatcacatgatagacttagagacaatacca
Db	2891	TTCCTTTGAATTGCGGGGTCTACTCCGAATTTACAAATGTAGACTTAGAGCAAAATACA
QY	2881	attctaatgcgttagcttggtagaagagctccacacacgcgaagggaaagagctttag
Db	2951	ATTCTAATGCGTTACTTGTGTAAGAGCTCCATACACTGCAAGGGAAGACCTTTTAA
QY	2941	attctgaagctcttcttgatctgatttcgaattccatttagttgttgcataaacatttc
Db	3011	ATTCTTGAAGGCTCTTTTCGATTTCCATTTGATTTAGTTGTGTAAGACACTTTC
RESULT	2	
LOCUS	CPN133035	
DEFINITION	CPN133035 26920 bp DNA BCT 24-FEB-1999	
ACCESSION	CPN133035	
VERSION	hbl, hb2, hb3, hb4, hb5, hb6, hb7, omp7, pseudogenes omp8, omp9, omp15, strain VR110.	
KEYWORDS	AJ133035.1 GI:4455890	
SOURCE	Hbl gene; HB2 gene; hb3 gene; hb4 gene; hb5 gene; hb6 gene; hb7 gene; omp15 gene; omp6 gene; omp7 gene; omp8 gene; omp9 gene; outere membrane protein 6; outer membrane protein 7; pseudogene.	
ORGANISM	Chlamydomophila pneumoniae	
REFERENCE	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.	
AUTHORS	1 (bases 1 to 26920)	
TITLE	Boesen, T.	
REFERENCE	Submitted (21-JAN-1999) Boesen T., Department of Medical Microbiology and Immunology, University of Aarhus, The Bartholin Building, DK-8000 Aarhus, DENMARK	
AUTHORS	2 (bases 1 to 26920)	
REFERENCE	Hjerno, K., Boesen, T., Daugaard, L., Knudsen, K., Madsen, A., Christiansen, G. and Birkelund, S.	
TITLE	Chlamydia proteins containing the GAT-repeat belong to a subfamily of autotransporting pathogenicity factors	
JOURNAL	Unpublished	
FEATURES		
SOURCE	Location/Qualifiers	
	1..26920	
	/organism="Chlamydomophila pneumoniae"	
	/strain="VR1310"	
	/db_xref="taxon:83558"	
	498..575	
	/gene="omp6"	
	498..3265	
	/gene="omp6"	
	/note="putative"	
	/codon_start=1	
	/transl_table=11	
	/product="outer membrane protein 6"	
	/protein_id="CAB37075.1"	
	/db_xref="GI:4455891"	
	/translation="MRSLSGCPPLVFSLTLLSVEDLSATITSLTPEDSPFGDSNA ERYINWADSVSLTGDVISISWNSALNKRCAFWTSQVTFAGNHHGLYFNVISGCT KRGAVLACQPOATARFSGSTLSFISQPEDIDEGCLSKRALMLANNYVREON QSTKRGALISGANVTIVGNDSVSYFONAAFFGALISSSPDLOIAYNWEIIFQNTA KNSGSGALTSDDIDIDQNNAYVLRRENDAULTATGKGAAYCCPLPSSGSSPVPITFS	

```

DNKQIVFERNHISMGGAIVARKLSISSGGPIFLINNITSYANSONLGAALIDTGGEL
SLSAKGTITFGNRTSLPFLNGIHLNNAFLKLOANNGSIEFYDITSEAGSTQ
LININGDPKNKEVGTILFSGEKLANDRDEFTIPQVNVLSAGVLVKEAEAVSTQ
LTQSPGSHLVLDIGTKLIASKEDAITGALIDISSSTAAVAKANTAKKOLISVT
SIELISPTGNVAFEDLRMNSOTFFLISLPGAGGSVTAGDFLPVSHYFOGQWKL
AMTCGNKVGEEFNKINYPKRPKEGVLVNNITMGNAVVRMSLOVETASSIQDT
RGLWIDIGNEFFHVASEDNIRYHNSGGVVLAVNNETPHYTSMAQSOLFSDKDY
AVSNNEYRMYLGSYLQYTTLSGIFRYASRPNVNVGILSRFLONPLMFHCAY
GHATNDKMTDYANPMPVKNWSMNMCWALIECGSMPLFQOGAIPPMKLOLVY
AYOGDFETADGRRFSNGSLTISVPIGFEFKALSDVLYDFSFSYIDIRKDP
SCEALVISGDSWLVPAHVSRHAFVSGTGTGHHNDITELLCKRSICRPHARYNI
NGSKFRP"
gene
498..3266
/ gene="omp6"
gene
3976..6870
/ gene="hb1"
3976..6870
/ gene="hb1"
/ note="hypothetical"
/ codon_start=1
/ transl_table=11
/ product="HB1 protein"
/ protein_id="CAB37076.1"
/ db_xref="GI:4455892"
/ translation="MRKRLAIVLAIIVASILLIAGCVLTVAIPLGSSVSPAGMG
ACAGCVLALGIDVLLKREVPVLAVITTPGSGPSGISISGASTIRSLPTYL
LDEGHPQSMRKLRLAIVLIVESILLIAGVLLVAIPGLSVISLSPAGACALGC
VMLAGIDVLLKREVPVLAVITTPGSGPSGISISGASTIRSLPTLPGHGP
OSMRKLRLAIVLIVESILLIAGVLLTVAIPLGSSVSPAGACALCVMLAG
IDVLLKREVPVLAVIPPEVVIDDIEDESLRLOGEAALALPPEMSAPCVIY
VESHLENNKSLPYOCHGLEKTKHDIRVRSKRAMVEPDIRIPPEEPFLSAR
KRDLATLTVKRLITLOLENNKRAFSYLIQDSIFKTIIDNEKIAMFLLNS
ICRFLITEHNEHGAKSLLKRNALVLEKVIYRSLOKSYDIGMSAMKILHNPFR
SLEDNKTKIMEHEMELSSRYKRLFDLSDENYDTPSPKMDLSGICRPLASE
ISRDEOMQRKAKHLQESLTYOARDRLDQSKENOKELKAEQOYISSMRVKEFI
ERVOERIAIOKLVNPIERBEETGOETVTPYOGTASSDLTDIGRITVSSRDS
ONOSCVAVLRSHEVMEVMEKOEYKPKKFOOMSGSLPFFTEHITELVLDKDS
KHLSEKRVNNKKEVOYAKFRKYLIESLBEILAOETESASLTOBELPILATGALS
KAVRGSLLCALASAKAPYFEDPFDSDTQALRLTLQEARASLEETKRSNME
NDLAEERLLKESKOTFERAGLIVLREIAYESTYDLRLTNTWGTPESEKVSMTL
NYNEKRRAKTRIVEMTORVRFMALEQMFNEALQELSLQAPSE"
7150..8058
/ gene="hb2"
7150..8058
/ gene="hb2"
/ note="hypothetical"
/ codon_start=1
/ transl_table=11
/ product="HB2 protein"
/ protein_id="CAB37077.1"
/ db_xref="GI:4455893"
/ translation="MAYFLSPQLPPPPRRSIGVCSKIRSLITLLVGLLTLGIP
GLTAGISGAGLGSALGVVLISGLLELVLRREVPYRSEIPRGVSVTSEEPALB
KAOKEPETKKILDLRPLKDLIDLTIVIQEVLACLEKEDRGLLEAKERLAVD
VVERDKMSEFIDIOVLNEBAYVYEHCCDPLENAYEIFFSSEOLRDYCAVGYLGS
GDARADRLKRSYKEVMDREKRVTKMSWASVMLDHSYVARELFRKAVGVLEESYKI
LEKSVDAFYCEKAKIORGRFKTL"
8185..9492
/ gene="hb3"
8185..9492
/ gene="hb3"
/ note="hypothetical"
/ codon_start=1
/ transl_table=11
/ product="HB3 protein"
/ protein_id="CAB37078.1"
/ db_xref="GI:4455894"
/ translation="MMLDRYADKFLIREKEEMRHEHFNATVYRKASGAYAKAKAA
FEKRSNENKRVKDVVEKWLKSLAEENHOSRRARERLRIOTLYPVSVEEVLAR
ORTKVNLENYADIERKYHCVABOEHYENKKEKYEIRENEKVLASAEVSECIO
RLBDCLETWSKKILTKAESVTEKFDATKGNKVLSDVNLRLILCEDAEMLFRIE
EIEMLTRVNEVPLLFMKNTFERKASLQVNSCKEMAKVPEOCKESPYRSSQERLERIN
QDLOTAYTNCORLQGSDESKEVTRCTCHDLREOMKHEVQGLNFINDEILLVGAELF

```

---

```

TOARLDIVATPYMEFYLOXHNIREKVRSGOMAKTEREIRIOAFQGVAKEDLLAD
TILKEEYWLRLDDWMLRBRKRNQRRLICNKILAAQQRVKGK"
gene
9805..10698
/ gene="hb4"
9805..10698
/ gene="hb4"
/ note="hypothetical"
/ codon_start=1
/ transl_table=11
/ product="HB4 protein"
/ protein_id="CAB37079.1"
/ db_xref="GI:4455895"
/ translation="MLDLISGALFTLGIPLGSLAASFGIGLISALGVLMISGLIC
LVKREIPVPEEIPPEGVSLAPSEEPALQAOETLQALPREIDOLDITDIOEVACIR
KIKDSYERSFLNDAKRELRFVDPEVVEDLSETEFLROIYVAOEGMINTLINGSRL
MTAESRSLDFHYSKRLGILPBGSDVYGBELKKSARKIYARLMSLGHETIKVYADR
NSYMAEKARAKALGALDESIVYNSLTOSYDKDFLESERAKIPMNGHITWLDDAKSQC
AEKRLGMPRVNVRNLGKQSGK"
10776..12119
/ gene="hb5"
10776..12119
/ gene="hb5"
/ note="hypothetical"
/ codon_start=1
/ transl_table=11
/ product="HB5 protein"
/ protein_id="CAB37080.1"
/ db_xref="GI:4455896"
/ translation="MDITTHLEKRTPIKRLHRKALAKTTEFKRSKKNQANVE
ANARRLVYDMDQOEOKAGEERLEKHALPEVSVIRNKKIQETRSNLEKAYALE
ENRCVREDEYDWEKREKRAERRENGNITLSEPELSELEQFDHGLKMFSEKMLL
EGHLKLOKQNTAEVENKILSDAESRLIEFYEDVKEMPCRIIEIEKTLRMAEPLPLPT
KKAERKQSYNSCAEMLEKVPYCKSLAVYSKELVSLDEDLRAAYECCORFQCG
DSGESRSLRREORIEPFOEGIDIVKELICVSSRLMTECCVSGVKEKAP
GKKRYAAYQEIYRVRVOSWMTMSSELRGVCANMKMLAAGLSEEDKYVKEEYWTY
REERKNEKRLVGTIKYIATQORIQEFPDSIVASSNKSVMKARFLNREDS"
12262..12987
/ gene="hb6"
12262..12987
/ gene="hb6"
/ note="hypothetical"
/ codon_start=1
/ transl_table=11
/ product="HB6 protein"
/ protein_id="CAB37081.1"
/ db_xref="GI:4455897"
/ translation="MTAPAILHVSPTPPETKFLVIPKDSKSRALGTLILVGLLIV
GALVYSGVSGISALIVCGIGISTISIGVLEVLGILLRRELTLDEIKQAIN
TPADELKELEMYIOSTEKSLEKIBGSRYSOGVLNRATORILDESLSITSEFNDL
RQLPDEKIELLSGERLLEFIANLFRQGDVYVNLGNLADIRAYGPNKYKVAWIE
KARAVHEFIVLTLMARELEFF"
13014..14585
/ gene="hb7"
13014..14585
/ gene="hb7"
/ note="hypothetical"
/ codon_start=1
/ transl_table=11
/ product="HB7 protein"
/ protein_id="CAB37082.1"
/ db_xref="GI:4455898"
/ translation="MTOENLRLERLLYNSVQSYADRLFSEYKTKMVDPTLPMEE
DKKCAEAERKALEOQRLILDYGSIFMLNDEINLNDPWSGLNTVTRKVFQOED
DSBRMNHKVLIOKLENDYENKLEESSEENKRLSLSDVLELDAKTAKFFLKQOE
VETRYKDLRARGYGVDPKQDTAKKKELELSEFLDSIESELVOCLEDDIYKKE
ODVKDLARQOELDEODIEAKREBAADIRSLNRLKRSKTMIDRAKHMENADSITYM
WTSQIEKMKAKRLKILKEIDTIVSLPEIDISLSEELPLTTTELTLKSTAKTI
CSRTILKMTVFNNTYVQDYETQOLNGLKGLGISOFRGKQDDPANNLEQVALCK
RLBELTONEFIQGFNMKEDEFKYAANDLYIRTAEOKNMFYQCMELFRYHHEVNRK
LLEMLNCAASYDARKKICSLRLDERELQKEIKKEEFYQKQOORHADRSRTTYOK
LRLAELELLELKKRI"
14954..17479
/ gene="omp7"

```

gene

CDS  
14954..17479  
/gene="omp/"  
/note="putative"  
/codon\_start=1  
/transl\_table=1  
/product="outer membrane protein 7"  
/protein\_id="CA37083.1"  
/db\_xref="GI:445899"  
/translation="MKIPLEFLISLVPLTSLNLSLGAATPELSNSFGDTSTTS  
FSKTSSTADGTVYKRVDSYVIEYVNPKEGTOSTSCFKMDAAGLNLGGGFSFTFS

Query Match 99.6%; Score 2988; DB 2; Length 26920;  
Best Local Similarity 100.0%; Pred. NO. 0;  
Matches 2999; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
QY 1 cgccttaaccagtagaggtgtgagtaattcctgaactgttcctccatgtgtatct 60
Db 17662 CGCTCTTACCTAGTAGAGGTGAGTGAATTTCTGACTTGTCTCTCTATGTGTAATCT 17721
QY 61 cttaaaatataatcaaatcaagtataattttaacatgaagtctcttccocaa 120
Db 17722 CTTAAATATTAAATCAAAATCAAGTATATTTTACAAAGAAAGTCTTCCCAA 17781
QY 121 gtttgatttctacattgtctcttccctgtctatgatgtctacgagacagtitt 180
Db 17782 GTTTGATTTTTCACATTTTGCTATTTTCCCTTGTCTATGATGTCACGACAGATTTT 17841
QY 181 ggaatcaagtgcaggttgcagtggaataaaatgttaatttcaagtcgtgagatca 240
Db 17842 GGATTCAGTGCAGATTGATGAGGAAATGGAATTTTTCAGTTCTGAGAGTCA 17901
QY 241 ggaagatgtcgaactacctaactttaaggaaatgtaactcagaataattccctgg 300
Db 17902 GGAAGATGTGGAACCTACCTATTAAAGGAATGTCACTAGAAATATTCTCG 17961
QY 301 aacagagcagcaatcacaaaagctgttttaacaacctaaaggcgattgaacttcaac 360
Db 17962 AACAGGCACAGCAATCACAAAAGCTGTTTACACACTAAAGGCGATTGACTTCAC 18021
QY 361 aggttaacgggaactctattgtttccaaacggtgatgcaggaactgaagagggctgc 420
Db 18022 AGGTACGGGAACCTCTATTGTTCCAAACGGTGATCAGAGACTGAGAGGGGCTGC 18081
QY 421 tgttaacagcagcgtgtgagataaattcaacacgtttataaggtttcttcgtatctt 480
Db 18082 TGTATACAGCAGCGGTGATGAATTAATCTACACGTTTATAGGGTTTCTTCGATCTT 18141
QY 481 tattgcgtctcctgggaagttcgataactacgcgcaaaaggacgttagctgcctcctgg 540
Db 18142 TATTGCGTCTCTGGAGTTTGATTACTACCGGCAAAAGGCCGTTAGCTGCTACGGG 18201
QY 541 tagcttgagtttgcacaaaatgtcaagtttgcctcttcagcaaaaactttcaacggataa 600
Db 18202 TAGCTTGAGTTGAC -AAAAATGTCAGTTTGCTTTCAGCAAAAACCTTTCAACGAGATA 18260
QY 601 tggcgtgtctatcacgcgcaaaaactcttcattaaacggagatacaatgtcagctgtt 660
Db 18261 TGGCGGTGCTATCACCGCAAAAACCTTTTATTAACAGGACTACAAATGTCAGCTTGTT 18320
QY 661 ttctgaaaataactccctcaaaaggaaggcgagccattcagactccgagtccttaacat 720
Db 18321 TTCTGAAAAATACCTCTCAAGAAAGCGGAGCCATTACAGACTTCGAGGCCCTTAACAT 18380
QY 721 tactgnaaaccaagggaagtcctctttcttgcataactcttcggattctgagctgc 780
Db 18381 TACTGNAACAACCAAGGGAAGTCTCTTTTCTGCAATTACTTCTTGAGATTCTGAGAGTC 18440
QY 781 aattttacgaagctcgtgtgactatttctaataatgctaaagttcccttaattgagaa 840
Db 18441 AATTTTTACGAAGCCTCGTGTACTTCTTAATAATGCTAAAGTTCTTATTATGACAA 18500
QY 841 taagtcacagagggagagctcctcaacaaacgggagatatgtcaggaggtgcatactg 900
|||||
```

```
Db 18501 TAAGTCAACAGAGGAGGAGCTCTCTCAACACGGGGGATATGTCAGGAGTCTATCTGTGC 18560
QY 901 ttataaaactagtaacagataagtaagtcacacctcaactggaatacagatgttactcttcag 960
Db 18561 TTATAAACTAGTACAGATTAAGTACCCCTACACGGAACACAGATGTACTCTTCAG 18620
QY 961 caacaatacatcgacaacacggcggagagctatcatatgtcaaaaagctcgaactgtctc 1020
Db 18621 CAACAATACATCGACAACACAGGGGAGGAGGATCTATCTATGTAATAAAGCTGAACTGGCTTC 18680
QY 1021 cggaggaattacccctattctcagtagaataagtgctcaatggaagtacagctccctaaagtg 1080
Db 18681 CGGAGGACTTACCTATTTCAAGTAATAGTGCATGAGGACACAGCTCTTAAGAGTGG 18740
QY 1081 agccatagctatcgaaatagtggaattgagtttatccgcgcgaatgaagtggaactgt 1140
Db 18741 AGCCATAGCTATCGAAGATAGTGGGAATTGAGTTTATCCGCGAATGAGTGGACATGT 18800
QY 1141 cttttggaataacagctcactctcactcctcctggacgaaataaagtagatcgaactt 1200
Db 18801 CTTTTTGGGAATACAGTACACTTCTACTCTCTGGGAGGAAATAGATATGACTT 18860
QY 1201 aggaacgagtgcaaaatgacagacttggtctcgtcgtgtagagccactactcta 1260
Db 18861 AGGAACGAGTGCAAAAGATGACAGCTTGGCTTCTGCTGCTGAGGACATCTACTCTTA 18920
QY 1261 tgatccataactcaagagatcatccaacaaagttaacagatgcttaaaagttaatgagac 1320
Db 18921 TGATCCATAACTACACGGGATCTCCACAAAGTTACAGATGCTTAAAGTATAGAC 18980
QY 1321 tccgagagattctgcactcaatatacatagaggaactacttcaacggagaagaattatc 1380
Db 18981 TCCGGAGATTCTGCACATCATATATACGGAACTCATCTTCACAGAGAAAAGTTATC 19040
QY 1381 agagaaagggccgcagagattcctaaacttacttcgaagtacatacagctgaaactt 1440
Db 19041 AGAGAACAGGGCGGAGATTTCAAAAATCTTACTCGAAGTACTTAACGCTGTACTCT 19100
QY 1441 ttcaagagtaactcattcttaaaaacatgagtgactctgcaagctcgaagctcaactca 1500
Db 19101 TTCAGAGAGTACTCTATTTTAAACATGAGTACCTGCAAGCTCAGGACATCTACTCA 19160
QY 1501 aagagcaatctcgtctcgaatggaactgtagaactcctgaagctcgtatctag 1560
Db 19161 ACAGGCAATCTCTGTCGAAATGAGACGTAGCACTGTAACCTGCTATCTAG 19220
QY 1561 caccataaacaatttggttcatttaacatcaagttctataagcgtgtgcaaaaggcaaaat 1620
Db 19221 CACCATTAACAATTGGTCATTAACTACATCAAGTTCTATAGAGGCAAAAGCAAAAAT 19280
QY 1621 agaaaccaaagctacgtcaaaaactcgaacttatctcgaacatcaactttaagacc 1680
Db 19281 AGAAACCAAAACCTACGTCAAAAATCTGACTTATCTGGAACATTACTTATGGACCC 19340
QY 1681 gaacggcagcttttaataaatacatagtttaagaactccctcagctcactagacatctaga 1740
Db 19341 GACGGGCAAGCTTTATGAAATCATAGTTTAAGAAATCTCTGACTCCTACGACATCTTGA 19400
QY 1741 gctcaaaagctctggaactgtaacaagcaccgcagtgactccaagatcctatattgggga 1800
Db 19401 GCTCAAAAGCTTCTGGAACGTAAACAGACCCGAGACTCCAGATCTCATATATGGGTGA 19460
QY 1801 gaaattccattagagctatcaaggaaacttgaggccaaattgttttggggagacagggcttc 1860
Db 19461 GAAATTCATTACGGCTATACAGGGAACCTTGGGCCCAATTGTTGGGGACAGGGGCTTC 19520
QY 1861 taagactgcaacctcaactggaactaaactgcatatttctaattcccgcgctatcgg 1920
Db 19521 TACAGCTGCAACCTCAATGTGACTTAAACAGTGGCTATATTCCTATATCCGAGCCTATCGG 19580
QY 1921 ctcttaagcccttaataagcttatggaatgcatltaagataatgactcctcactatct 1980
Db 19581 CTCTTAAGTCCCTAATAGCTTATGGAATGCATTTATAGATTTAGCTCTCTCCATTTCT 19640
```

QY 1981 tatggaagctcaacagaggggttcagggagaccgtcttttggctgtgatac 2040  
 |||||  
 Db 19641 TATGAGAGCTCAACGAGAGGGTTCCAGGAGACCGTCTTTTGGTGGGATTATC 19700  
 |||||  
 QY 2041 taactcttcataagagtagtacaaacacagcgcgggttcgcacattgaatgacg 2100  
 |||||  
 Db 19701 TAACTTCTTCATAGAGTAGAGTACAAAACACGAGCGGGGTTCCCATTTGAGGGCGG 19760  
 |||||  
 QY 2101 ttaatgcaagaggaacaccataactgttcagataagattcttagtgcatttg 2160  
 |||||  
 Db 19761 TTAATCTCATTAGAGAGAACTACATCTGTTACAGATAGATTCTTAAGTGGCTGCTTTG 19820  
 |||||  
 QY 2161 tcaagctcttggagaagataagactacttggtagtaagaatcaaggtaagctacag 2220  
 |||||  
 Db 19821 TCAGCTCTTGGAGAGATAGAGACTACTTGTAGCTAAGATCAAGGTACAGTCTACGG 19880  
 |||||  
 QY 2221 aggaactctcattaccacgacaaagaacctatactctcttccttcgcaactacggcc 2280  
 |||||  
 Db 19881 AGGAACCTCTATTACACGACACAAAGAAACCTATATCTCTCTCCATCAACACTACGGCC 19940  
 |||||  
 QY 2281 ttgttcgttgccttaattgtcttcacagagattccgttctcttcttcagaaacctagcta 2340  
 |||||  
 Db 19941 TTGTTCCGTTGCTTATGTTCTCTACAGAAATTCCTTCTCTTTTCAAGAAACCTTAGCTA 20000  
 |||||  
 QY 2341 caccatacgcgataacgaactcgaanaaccagltatacaacatactcactgttaagaag 2400  
 |||||  
 Db 20001 CACCCATACGGATACGATCTGAAAACCAAGTATACACATATCTACTGTTAAAGNAG 20060  
 |||||  
 QY 2401 ctggggagatagattcgcgttttagaattcgggtggagaagctcgcatttcttagatga 2460  
 |||||  
 Db 20061 CTGGGGGAATATAGATTTCGCTTTAGAAATCGGAGTGGAGAGCTCGAATTTGCTTAGATGA 20120  
 |||||  
 QY 2461 aagtgctcattttagcagcatatgccttcatagaattcagttgtctatgcacatca 2520  
 |||||  
 Db 20121 AAGTGCTCTATTGAGCAGTACATGCCCTCATGAATATGAGTTGCTATGACATCA 20180  
 |||||  
 QY 2521 ggaaggtttaaagaacagggaaacgaagctcgtgaaatttggaaagtaagcgtctgtgaa 2580  
 |||||  
 Db 20181 GGAAGGTTTAAAGAACGGGAAACGAAAGCTGTAATTTGAAATACCCCTCTTGTGAA 20240  
 |||||  
 QY 2581 tcttgccctaccatacgcgataccgatttgataagaagaacagctccagatagcacgta 2640  
 |||||  
 Db 20241 TCTTGCCCTTACTATCGGGATCCGATTGATAGGAATCAAGCTGCCAAGATGCAACGTA 20300  
 |||||  
 QY 2641 caatcctccttggtatctatcgtgagatctgttcgtatgaaccgcgactgtaacgacac 2700  
 |||||  
 Db 20301 CAATCTACTCTTGCTATCTGATGATCTGTTGCTGATGTAACCCGACTGACGACAAAC 20360  
 |||||  
 QY 2701 actgcgaattgagcggtgttcttggaaaacctcgtgtaacgaatttggcaagaacagctt 2760  
 |||||  
 Db 20361 ACTGGGAATTAGCGGTGATTTTGGAAAAACCTTGGATTCGAAATTTGGCAAAACAGCTTT 20420  
 |||||  
 QY 2761 agctcctcgtcagggaaacacatttcttcttaactcaaatcttgaagccttagcaatt 2820  
 |||||  
 Db 20421 AGTCCTTGTCAGAGGACCAATTTTGTCTTAACTCAATTTTGAAGCTTTAGCCAAAT 20480  
 |||||  
 QY 2821 ttccttgaattcgtgggttcactcgcgaattacaaatgtaacttaagagcaaatataca 2880  
 |||||  
 Db 20481 TTTCTTTGAATTCGGTGGGTATCTCCCAATTTACATGTATGAGAGCAAAATACCA 20540  
 |||||  
 QY 2881 attcctaaagctttagcttggtaaagagctcatacatcgaagggaagaagcctttaaag 2940  
 |||||  
 Db 20541 ATTCTAATGCTTACCTTTGGTAAAGAGCTCATACATCGAAGGAAAGAGCTTTAAG 20600  
 |||||  
 QY 2941 attcttgaagcctcttctcgaattcgaattcgaatttagtgatttgcataaacacttc 3000  
 |||||  
 Db 20601 ATTCTTGAAGGCTCTTTGCATTTGCATTTCAATTTAGAGTTTGTCTAAACACTTTC 20660  
 |||||

RESULT 3  
 AE002235/c AE002235 12127 bp DNA BCT 30-MAY-2000  
 LOCUS

DEFINITION Chlamydomophila pneumoniae AR39, section 63 of 94 of the complete genome.  
 ACCESSION AE002235  
 VERSION AE002235.2  
 KEYWORDS GI:8163495  
 ORGANISM Chlamydomophila pneumoniae AR39.  
 Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.  
 REFERENCE  
 AUTHORS Read, T.D., Brunham, R.C., Shen, C., Gill, S.R., Heidelberg, J.F., White, O., Hickey, E.K., Peterson, J., Unayam, L.A., Uitterback, T., Berry, K., Bass, S., Linher, K., Weidman, J., Khouri, H., Craven, B., Bowman, C., Dodson, R., Gwin, M., Nelson, W., Deboy, R., Kolonay, J., McClarty, G., Salzberg, S.L., Eisen, J., and Fraser, C.M.  
 TITLE Genome sequences of Chlamydia trachomatis Mohn and Chlamydia pneumoniae AR39  
 JOURNAL Nucleic Acids Res. 28 (6), 1397-1406 (2000)  
 MEDLINE 20150255  
 REFERENCE 2 (bases 1 to 12127)  
 AUTHORS Read, T.D., Brunham, R.C., Shen, C., Gill, S.R., Heidelberg, J.F., White, O., Hickey, E.K., Peterson, J., Unayam, L.A., Uitterback, T., Berry, K., Bass, S., Linher, K., Weidman, J., Khouri, H., Craven, B., Bowman, C., Dodson, R., Gwin, M., Nelson, W., Deboy, R., Kolonay, J., McClarty, G., Salzberg, S.L., Eisen, J., and Fraser, C.M.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA  
 COMMENT On Jun 1, 2000 this sequence version replaced gi:7189672.  
 FEATURES  
 source location/Qualifiers  
 1..12127  
 /organism="Chlamydomophila pneumoniae AR39"  
 /strain="AR39"  
 /db\_xref="taxon:115711"  
 /note="synonym: Chlamydia pneumoniae AR39"  
 complement(150..2987)  
 /gene="CP0757"  
 /note="This region contains an authentic point mutation, causing a premature stop, and is not the result of a sequencing artifact; similar to GP:4455886, identified by sequence similarity; putative:polymorphic membrane protein G family authentic point mutation"  
 complement(3003..3191)  
 /gene="CP0758"  
 complement(3003..3191)  
 /note="hypothetical protein; identified by Glimmer2; putative"  
 /codon\_start=1  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="AF38560.1"  
 /db\_xref="GI:7189673"  
 /strand="+"  
 /translation="MRLLSCPMLFRAACAFSGFQEMQGRNIOSLDMMASLSLELFSTRKAGVSLSELRHRTAR"  
 complement(3336..6086)  
 /gene="CP0759"  
 /note="This region contains a gene with one or more premature stops or frameshifts, and is not the result of a sequencing artifact; similar to GP:4455886, identified by sequence similarity; putative:polymorphic membrane protein G family degenerate"  
 complement(6125..9110)  
 /gene="CP0760"  
 /note="This region contains an authentic frame shift and is not the result of a sequencing artifact; similar to GP:4455886, identified by sequence similarity; putative:polymorphic membrane protein G family, authentic frame shift"  
 complement(9393..11918)  
 /gene="CP0761"  
 complement(9393..11918)  
 /gene="CP0761"  
 /note="similar to GP:4376270; identified by sequence

```

similarity: putative"
/codon_start=1
/transl_table=11
/product="polymorphic membrane protein G family"
/protein_id="AA038561.1"
/db_xref="GI:7189674"
/translation="MKIPFLILSLPTLSMNSLGAATTELSASNSPDGTTSTG
FSKSTSSARDGTVNYVFKDSVTLENPKRGTEGTSTGCFKNDAAAGLNLGGFSTFE
INDATTAASGAIGSSBAARATVTLNLSFSAISFKSPASTVTVNGAIINVKGNLSLDLNT
KLIQDNPEPTAGCGGAIINCAQSLKIANNSLSTIGSSSTRGAIHITKMLTSSGEGTTP
FGGNAPRTAAGKGAIALADSGTSLISDQSGIILEGNTGIVTGHVSDIGLGSAA
ITALAAGQHTIPEYDPITVTGSTVADLANSPDSTDNKEGTVTFYSEKULTEAR
ADEKRNKSTKLQNAVFNRGTVLKGVDVLSANGSPOAANCLIMDLDTSLVANTES
LEFLNEINISLNRNKRKIKLSAATKORIDRIDRPIAISPSTQNFQLEEDSSTYAE
ILELDAGKDIVISADRSIDAVQSPYQSGQKWTINWSTDKKAVISAKQSFNPAQEC
EAPLVPNLNLSGSDIVRSFQNFIELGTAIEYERFWAGVLSNVLHRSRGKQKRRFRR
VSGAGVVGASTRMGQDLSLGFQLEARDKDYEMNINFAATVAGSLRLQHDASLVS
VSLILGEGSLRILPYSKTLPCSFSTQSLGYGTHDRMKTESLPPPTLTDTHTSM
GGYVWAGELGTRVAVENTSGRGFOEYTPFAKVOAYVARODSFVEIGALISRDEPSHL
INLAIPLGIEKRFPAEQLYHVAVMAYSDVCVRSKPTTTLISQSSWKTIGSLNARQ
AGIVQASQFSRLGAAALEFGNFGWEWRSSSYVNDAGSKTF"

```

Query Match	99.58;	Score 2986.4;	DB 1;	Length 12127;
Best Local Similarity	99.98;	Pred. No. 0;		
Matches 2998; Conservative	0;	Mismatches 1;		

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

QY	661	ttctgaaaaatacctctctcaagaagcgagccattccagctccgattgcccctacat	720
Db	8551	TTCTGAAAAATACCTCTCTCAAGAAAGGCGAGCATTCAGACTTCGATGCCCTTACCAT	84922
QY	721	tactggaacccaaggggaaagtcctctttcttcaaacatctcttcgaalettcgaaetgc	780
Db	8491	TACTGGAACCAAGGGGAAGTCTCTTTCTCTACAAATCTTCTTGATTCGAGCTGC	84322
QY	781	aattttacgaaagcctcggtgactatttctataatgctaagafttcccttatgcaaa	840
Db	8431	AAATTTTACGAAGCCCTCGTGACTATTTCATATAATCTAAAGTTTCCTTATTGACA	8372
QY	841	taagtcacaagcgagctcctctcaaacagggggatctgctcagaaggtgctactgctg	900
Db	8371	TAAGTCAACAGAGGAGAGCTCCTCAACAAACGGGGATATGTCAAGAGGTGCTATCTGTGC	8312
QY	901	ttataaacttagtacagatactaaagtgtaacctcacttgaaatagatggtactcttag	960
Db	8311	TTATAAAACTAGTACAGATACTAAGTAGTCAACCTCAGTCGAAATAGATGTTACTCTTCAG	8252
QY	961	caacaatatactgacacacagggggagagctcattatgtaaaaaagctcgaactgcttc	1020
Db	8251	CAACAATATATGACACACAGCGGGAGGAGCTATGTATGAAAAAGCTCCAACTGGCTTC	8192
QY	1021	cgaggaacttaccctattcagtaagaatagtgtaatgagaggtgacagctcttaagtg	1080
Db	8191	CGAGAGACTTACCCATTCACATAGAAATAGTGCATAGAGGTGACAGCTCTTAAAGTGG	8132
QY	1081	agccatagctatcgaaagtagtgggaaattgaattatccgcgatagtggtgatactgt	1140
Db	8131	AGCCATACTCTCCAAAGATATGTGGGAATTGAATTATCCGCCGATAGTGGTACATGT	8072
QY	1141	cttttaaggaaatacagtcactcttactactccctggagcgaataagaagratgact	1200
Db	8071	CTTTTAAAGGAATACAGTCACTCTCTACTACTCCTGGAGAGATAGAGTAGTATGACT	8012
QY	1201	aggacgaagtccaagaatgacagcttggcttctgctgctgtgtagagccatctactta	1260
Db	8011	AGGAACGAGTGCAAAGATGACACACTTGGCTGTGCTGCTGATAGGCATCTACTTCTA	7952
QY	1261	tgtaccataactacagagatcatccacaacagttacagatgtgcttaaaagttaagagac	1320
Db	7951	TGAATCCCATACTACAGAGATCATCCACACAGTTACAGATGTCTTAAAGTTAATGAAAC	7892
QY	1321	ttccgacagatctcgacactcaatatatcacggaaatcattcacaggaaagaagtatc	1380
Db	7891	TCCGGCAGATTCGCACTCAATATATACAGGAACATCATCTTCACAGGAAAGATTATC	7832
QY	1381	agagacagagcgccgacatctcacaatatcttaactctcgaaagtactaaagcctgtaactc	1440
Db	7831	AGAGACAGAGCGCGCGAGATTCTTAAAAATCTTAAGTCTGAAAGCTAATCAGCCTGTAACTCT	7772
QY	1441	tttaagaggtactctatcttttaaacatgagagtgactctgacatcgaacttactca	1500
Db	7771	TTTCAAGAGTACTCTATCTTTTAAACAAGAGATCACTCTGAGACTCAGGATCTCACTCA	7712
QY	1501	acaagcagattctgctctcgaaatgagcgtgaggaactactctgaacactgctatactag	1560
Db	7711	ACAGGCAGATTCCTGCTCGAAATGAGACTGGAATCACTTCAAACTCTGTATCTATA	7652
QY	1561	ccccaataaacaattggtcatataacatcagttctataagacgtgtaacagaagggcaaaat	1620
Db	7651	CACCAATTAACATTTGGTATTAACATCAAGTCTTAATAGAGGCTGCAAGAGCAAAAT	7592
QY	1621	agaaaccaaagctacgtcaaaaaatctgacttatactcggaacataactatttgaacc	1680
Db	7591	AGAAACCAAAGCTACGTCAAAAAATCTGACTTATCTCGAAACCATCACTTATGACCC	7532
QY	1681	gacgggcacgttttatgaaaaatcatagtttaagaaatcctcagttctcagaactctaga	1740
Db	7531	GACGGCACGTTTATGAAAAATCATATCTTAAAGAAATCCCACTCTCTAGACATCTTGA	7472
QY	1741	gtctaaagcttctgactcttaacaagcaccgagtgactccaagatctctataatggtgta	1800



```

|||||
Db 7471 GCTCAAGCTTCTGAACTGTAACAGCACCAGTACCTCCAGATCCATATATGGGTGA 7412
Oy 1801 gaattccattacggctatcagggaacttggggcccaattgtttgggggaacggggttc 1860
Db 7411 GAAATTTCCATTACGGCTATCAGGAACTTGGGGCCCAATTTGTTGGGGACAGGGGCTTC 7352
Oy 1861 taagactgcaaccttcaacttgaactaaactgtctatcttccaaatcccgagcgtatcgg 1920
Db 7351 TACGATGCAACTTCACTCACTGAGCTAAACTGGCTATATTCCTAATCCGACGATATCGG 7292
Oy 1921 ccttttagtcccttaataactttatggaaatgcatattatagatctccattatct 1980
Db 7291 CTTTATAGTCCCTAATAGCTTATGGAATCCATTATAGATATAGCTCTCTCCATATCT 7232
Oy 1981 tatggaactgcaaaaggaagggtttgaggagagaccgtgtctttgttggtgcgtatc 2040
Db 7231 TATGGAGACTGCAGAAAGAGAGGTTCGAGGAGACCGTCTTTTGTGCTGCTGATATAC 7172
Oy 2041 taactcttccataaggaatagatcaaaaacacgacgagggtttcgccattgagtgcg 2100
Db 7171 TAACCTCTCCCAATAGGATAGTACAAAAACAGCAGCGGGTTCGCCATTGTGAGTGGCGG 7112
Oy 2101 ttaattcaaggaggaagaacctatatactgttcagataagatcttaagtgctgcatttg 2160
Db 7111 TATGTCATAGAGAGAAACCTACATCTGTTCTCAGATTAAGATTTAGGCTGCAATTTTG 7052
Oy 2161 tcaagctctttgaaagagatagactactcttgactaaagaaagacagctcagc 2220
Db 7051 TCAGCTCTTGGAGAGATAGAGACTTGTAGCTAAGATCAAGTACAGTCTACGG 6992
Oy 2221 aggaactctatctaacagacacgaacactatctcttctccttggaaaactacggcc 2280
Db 6991 AGGAACCTCTCTATACACACACAGCAAACTATATCTCTCTCTGCAAACTACGGCC 6932
Oy 2281 ttgttcgtgtcttcttcttcacagagattctgtctccttttaaggaaactatgcta 2340
Db 6931 TTGTTGCTGTCTTATGTCTTACAGAGATCTGTTCTCTTTCAGGAACCTTAGCTA 6872
Oy 2341 caccatacagataacgacttgaacacacgaagatacaacatactactcgttaaaagaa 2400
Db 6871 CACCCATACGAGTAACGATCTGAAAAACAAGTATACAACTATCCACTGTTAAAGAAAG 6812
Oy 2401 ctgggggaattagttctgcgtttagaattcgggtggaagagctccgattgcttagata 2460
Db 6811 CTGGGGGAATGATGTTGCTTGAATTCGTTGAAAGCTCGGATTTGCTTAGATGA 6752
Oy 2461 aagtcctcatttgaagcagatcatgcctctcagaattcgaagttgtctatgacatca 2520
Db 6751 AAGTCTCTATTGAGCACTACATGCCCTTCTGAAATTCGATTTGCTTATGCAATCA 6692
Oy 2521 ggaaggtcttaagaacaggaagcgcgtgaatttggaaatgacgcgtctgtgaa 2580
Db 6691 GGAAGGTTTAAAGAACAGGGAAGAAAGCTGTAATTTGGAAGTGAAGCCGCTTGTGA 6632
Oy 2581 tcttgccttaccatcgggagtcggatttgataaggaatgagctgcgaagatgcaacgta 2640
Db 6631 TTTTGCCTTACCTATCGGATCGGATTTGATTAAGGATAGATGCAAGATGCAACGTA 6572
Oy 2641 caatctaactcttggttatactgtgacatctgtctgtaglaaccccgactgtaacaaac 2700
Db 6571 CAATCTAATCTTGGTTATACGTGGATCTTGTGTAATACCCGAGTATACGACAAAC 6512
Oy 2701 actgcaatctagcgtgacatcttggaaaaactcggtagcgaatttggcaagaagcttt 2760
Db 6511 ACTGGAAATTAACGGGTGATTTTGGAAAACTTTCGTAAGATTTGGCAAGTGAAGCTTT 6452
Oy 2761 agtcccttgcagggaacacatttggctttaactcaaattttgaagccttaagccaatt 2820
Db 6451 AGTCTCTCTGAGGGAACCATTTTTCCTTAACTCAAAATTTTGAAGCCTTAGCCAAAT 6392
Oy 2821 tctcttgaattgcgtgggtcatctcgaatctatagacttaggaagcaaaatacca 2880
|||||

```

```

Db 6391 TTCTTTGAATTCGCTGGGTCATCTCCGAATTACATGTAGACTTAGAGCAAAATACCA 6332
Oy 2881 attcctaatgcgttagccttggtaaagagctcatatcatcgaaggaagaaagcctttag 2940
Db 6331 ATTCTTAATGCGCTTACTTGGTAAAGAGCTCCATACATGAGGAAAAAGAGCTTTTAG 6272
Oy 2941 attcttaagagcctcttcgacttgcattccatcttccattagtgcttggctaaacacttc 3000
Db 6271 ATTCTTGAAGGCTTTTTCGATTTTCGATTTTCATTTTAGTGTGTTTGTCTAAACACTTC 6212

RESULT 4
AP002545
LOCUS
DEFINITION
Chlamydomonada pneumoniae genomic DNA, complete sequence, section 1/4.
ACCESSION
AP002545
VERSION
AP002545.1 GI:8978373
KEYWORDS
SOURCE
ORGANISM
Chlamydomonada pneumoniae (strain:J138) DNA.
REFERENCE
Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomonada.
AUTHORS
Shirai, M., Hirakawa, H., Kimoto, M., Tabuchi, M., Kishi, F., Ouchi, K., Shirai, T., Ishii, K., Hattori, M., Kuhara, S., and Nakazawa, T.
TITLE
Comparison of whole genome sequences of chlamydia pneumoniae J138 from Japan and CWD29 from USA
JOURNAL
Nucleic Acids Res. 28 (12), 2311-2314 (2000)
MEDLINE
20330349
REFERENCE
2 (bases 1 to 300050)
AUTHORS
Shirai, M.
TITLE
Direct Submission
JOURNAL
Submitted (04-JUL-2000) to the DDBJ/EMBL/GenBank databases.
Mutsunori Shirai, Yamaguchi University School of Medicine, Department of Microbiology; 1-1-1 Minamikuogushi, Ube, Yamaguchi 755-8505, Japan (E-mail: mshirai@med.u-u.ac.jp, Tel.: 81-836-22-2227, Fax: 81-836-22-2415)
COMMENT
AB033780-AB033781, AB033792-AB033799: Submitted (25-Oct-1999)
AB038345-AB038347: Submitted (14-Feb-2000)
AB036071-AB036078: Submitted (18-Dec-2000).
Location/Qualifiers
1..300050
/organism="Chlamydomonada pneumoniae"
/strain="J138"
/db_xref="taxon:83558"
/note="synonym: Chlamydia pneumoniae"
complement(1..282)
/gene="CpJ0001"
complement(1..282)
/gene="CpJ0001"
/codon_start=1
/transl_table=11
/product="CT001 hypothetical protein"
/protein_id="BA08211.1"
/db_xref="GI:8978373"
/translation="MLGKIRIGLSLIVICALNVGLIGTHNKLNIAKTCGVSP
AQTQYIIRIAGVYICLSFPCSKSRSHSDSCSGCHSHSDKN"
573..878
/gene="gacA"
573..878
/gene="gacA"
/codon_start=1
/transl_table=11
/product="Glu-tRNA Gln amidotransferase (C subunit)"
/protein_id="BA08212.1"
/db_xref="GI:8978373"
/translation="MEQFLDRREIILLAKASALQISELDIOYQSLSAVITSMKEA
LAIEIDPADSCESLFMHVNVEDLREDSVTSDFNREBFLRVNPEISGLGVKPAVAK"
895..2373
/gene="gataA"
gene

```

CDS  
895..2373  
/gene="gata"  
/codon\_start=1  
/transl\_table=11  
/product="Glu tRNA Gln amidotransferase (A subunit)"  
/protein\_id="BA98213.1"  
/db\_xref="GI:8978376"  
/translation="MIRTSALDELAKAVTIGELTANGVYQHFFHRIIEAEQVGAFTSL  
CEQALEQALIDKRRSGPELKLAVGVPVGIKDNIVHTGLTKTCASRVLENTQPPED  
ATVPRIRIKEDGIIIGKLPMDDEPMAGSTLLYSAPFTANPMDLSKVPQSGSGGSAAY  
SARFCPVAGSDTGSGIRQPAFCGVGGRKPSGAVSRXGLVAFASLIDQIGPLANTV  
EDVALMVDVFGSDPDKDSTREFRDSFMSKLTSTVPKYIGVPRTEGLADIDREVF  
FSSLAIFEGETHLVDELIDILSHAVSYIIASAATNLARFQVGLGYRSGQVAT  
ISQVLDLRSRGGFKGVRRLLGNVYLSEARONVYKATATVAKATVAKRATFEFC  
ETLAMPVCSAPAEIGETLVTLVLYODIYVAMNLAATPAIAYSGSKSGDPLGLG  
ILGQGGDOGVCOYGVSGFQEHQAIKQLFSKRAKASVYLGGGS"  
2370..3836  
/gene="gataB"  
2370..3836  
/gene="gataB"  
/codon\_start=1  
/transl\_table=11  
/product="(Pec112) Glu tRNA Gln amidotransferase (B  
subunit)"  
/protein\_id="BA98214.1"  
/db\_xref="GI:8978377"  
/translation="MSAVYADMSGVIGLEVHETANASKLFSSALNRPGRDPNTIST  
VCTGLPSGLPVINOSAVKAVLFGCAVEGSIILSPDKSYFDPSPRNQITQFPH  
PIIRGRIRKAIYQGESEYFELAQHIEDDAGMKHGFEGACVDYRACVPLTEVSR  
CMFCPEDAVYATISVSLDYIGISDCNMEGSTRDVAVSRPKSFELNKEIKR  
MNSFEMQALAEKQRIIDTLMQPKDKLTIPATIRMDPEKKTVALRLESKE  
DKYFPEPDLPLQLTESYIERIKTLPETPKYHRYIOEGSEDAIILSDKNI  
AFPEVACDKCKNFRSLSNWVSEFGRCCTLVGKIPSSGIPPEGVQLVAIDQVY  
TKIKAIKELADLMSPGKNEPEIKKEPELIPMSDEGLQKTIARVLANPESIVDK  
NGTKALGLVGOIMKRTAGKAPPRVNEILLIDKNG"  
4127..6895  
/gene="pmp\_1"  
4127..6895  
/gene="pmp\_1"  
/codon\_start=1  
/transl\_table=11  
/product="polymorphic outer membrane protein G family"  
/protein\_id="BA98215.1"  
/db\_xref="GI:8978378"  
/translation="WRFSLCGFPIVSEFTLLSVFTSLSATTISLTPEDSFGDSQNA  
ERSYVQAGDVSLTGDVSLSNVDSNLNACPNVTSGSVTFAGNHGLNPNNTISGT  
TREGAVLCCQDPQATARSGETLSFTQSGDIKEGCLYSKNALMLNPNVREFEON  
OSRTKGGAISQANVTIVGNDSVSYFQNAITFGAISHSGVCLPTSGSSTVPVITGS  
KNGSGALYSDGIDIDQNAVLFRENEALTTAIGKGAVCCLPTSGSSTVPVITGS  
DKQALVERHNSIMGGAIVARKLISSGPTLFINNISANSOVLGAIAIDTGEI  
SLSAEGTITFGNRTSLPPLNGIHLQNAKFLQARNGSYIEFYDITSEADGSTO  
LNINGDPKRRKTYGTILSEGKSLANDPDKETIPQNVNLSAGYLVKEGAETVSK  
FVQSGSHVLDLGTKLIAKSEDAITGLAIDISLSSSTAIVAKATKAKOISVND  
STELISPTGNAYEDLRMNSQTFPALTLEPGAGSVTAGDPLPSPVHGFQGNMML  
AWMTGKNVGEFEFMDKITVXRPRKEGNTVNTLNGNAVDVRSILMOVQETASLQD  
RLMLTIDIGNPFRHASSEDNRVHNSGCVLYSVNNEITPKHTMAHSQLPSKDKY  
AVSNNEPRATLSTLIQITTSIGMTIRIARKNVAVGILSKRFLQNLMLFHLKAY  
GHATIDMTDIANPEPMVNSKRNCAIEBGGSPLVLENGRLQGAIPMKIQLVY  
ATQGDPKETIADGRFNSGSLTISVPLIGRFLKALSDVLYDFSFSYIDIRKQD  
SCBALVYISGDSWLPAAHVSRRHAFVSGRGVHFNDYELLCRGSIIECRPHANNYI  
NCGSKFRF"  
complement(7138..7293)  
/gene="CP10006"  
complement(7138..7293)  
/gene="CP10006"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BA98216.1"  
/db\_xref="GI:8978379"  
/translation="MQEPLRSALDERLSEWLVLLGVSPETTRSTPEKANQLPKDSR  
NRLTESL"  
7605..10499  
gene

CDS  
/gene="CP10007"  
7605..10499  
/gene="CP10007"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BA98217.1"  
/db\_xref="GI:8978380"  
/translation="WRKIRLIALVLAISLIIAGGVLLITVAIPGLSSVTSBPAGNG  
AALCCQALGIDVLTKKREVPVLAAYTPTEGVSISGASTIRSLPTL  
LDEGPOSKRKIRLIALVLYFSIIIASGVLLITVAISGSSVTSBPARGACALCC  
VMAIGVILVKKREVPVLAAYTPTEGVSISGASTIRSLPTLDEGHP  
QSKRKIRLIALVLYFSIIIASGVLLITVAIPGLSSVTSBPARGACALCCVLAIG  
IDVLKKREVPVLAAYTPTEGVSISGASTIRSLPTLDEGHP  
VSHLENNKSLPYDGHGLEKTKHQRVAVSSILKAVPEFLIRIFEEFEFLYSR  
KLIIDLATTVKREKTLTDEGLRNNKRAVSYLQDSIFKIIIDNEKIAMKEMTLSS  
ICRFTIIFENHGVAKSLIKRNVLLVAVYRSLOKSYRDIQNSATKTIHQNP  
SLEDNKTIMKEHAMELSYKATFLASDENVYDTPSDKKNDSIGTCRNLSE  
ISDEWQKAKLKHQESLITQAKRDITDQSSKRNKELKAEQRTISSRVRKFEI  
ERYQERIQALQILPNIILREBETGQVTPVQGTLSLSQTLIDLRILEVSRDN  
QNOESCVALRSHEVMSWEVKQEGPKRKEFQDQSGSLERFTEHIBELVLOKDY  
KHLSTFRVNNKKEVOYAKRKLKYLESDGLAQTESXESLITQEDLPILATGALC  
KAVFGSGLCCALASAKAPYEDPREFQDSTQIALTLRLQAKSLDEELKRSNLE  
NIAEBERRLAKSKQTEFRAGLGVLEIAVESYIDRLSTNWEETPSESEVYSMTL  
NYNEERKAKRTIEMTQRYRDEKMALEMCNENBALQEBLSIQARSE"  
10975..11688  
/gene="CP10008"  
10975..11688  
/gene="CP10008"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BA98218.1"  
/db\_xref="GI:8978381"  
/translation="MISGLFLVLRPEVPTVRSEELPRGVSVTPSEEPALKAQKE  
TKILDRLEKELDQDITRYIOEVFACLERLDEPDRGLITRAKELRFPVVERDKM  
SEFLDIOVRLNEAYVYEHQDPLENIAVYIFSSGLRQYACGCVLPGSDRAAR  
LRSYKVEVMDRMRYTKSMWASVLDHSGVARELFRKAVGLVEESYIKILFESYD  
AFKECEKAKIQRDGRFKWL"  
11815..13122  
/gene="CP10009"  
11815..13122  
/gene="CP10009"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BA98219.1"  
/db\_xref="GI:8978382"  
/translation="MWLDRYDKFILREKEKMERHELPHATMVKKAAGVAKAKAA  
FEKERENENORVKVEKWLKSGLAEPNOSRRARERLRELOTLYPEVSEERLLE  
ORFKRYNLNIAADIEKTYHCVQEOHYKREVENEARVRENGKVAIAEBSFQ  
RLDDCEMTSKSLTKAESEVEMKEDTEGLKAVLSDVTNLEILICDAEMFRIE  
EIMETIRAMEVLEPLFMKTFERKASIOVNSKEMLAVEQCESTYISSEELRPLN  
ODIOTVYTCORLOGESDLRSKRYCTORLROMHFPVOCALININELIAYCAEFL  
TQNRIDLVATVYVMEFYIQYINIKREVRVQMAKTERRELRQAFQVYKEDLIAED  
TILKEDVILWLDMLDERKNRRLICNKIAAQOKVKG"  
13434..15758  
/gene="CP10010"  
13434..15758  
/gene="CP10010"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="BA98220.1"  
/db\_xref="GI:8978383"  
/translation="MLLISGALFLTIGTIGLSAASIFGLGSLNAGVLMISGLIC  
LTVKRIIPVRPEELPBGVSLAPSEPALDAAKTTIAQPKLEIDQDLDIDQVACLR

Query Match 99.5%; Score 2986.4; DB 1; Length 300050;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2998; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 1 cgccttaccagtagaggtgtagtaattcttgaactgtttctcattatggtgatatct 60

Db	21300	CGCTTTACCTGCTGAGAGGTTGAGTGAATTTCTTGACTGTTCTCCCAATGGTGATAT	21359
Qy	61	cttcaaatatfatatcatcaaatcaaatatataattttacaatgaagctctcttcccaca	120
Db	21360	CTTAAATATTAAATTCAAAACCAAGATATATATTTTACAAAGAAGTCCTCTTCCCA	21419
Qy	121	gttttatcttcacatttgcattcttccctttgtctaatgtctgaattgtctacagaaagttt	180
Db	21420	GTTTGTATTTTTCTACATTTGGCATTTTCCCTTTGCTCATGATTTGCTACGAAACAGTTT	21479
Qy	181	ggattcaagctcgagtttcgatlyggaataaaatgyltaatttttcaagtcgtgagatca	240
Db	21480	GGATTCAAGTGGAGATTTCGATGGGAATTAATAATGGTATTTTTCAGTTCCGGAAGTCA	21539
Qy	241	ggaagatctgtgaactaccctacctaattbaaggaaatgcactctagaataattctctg	300
Db	21540	GGAAATGCTGGAACTACTACTCTTTTAAAGGAATGTCACCTCTAGAAATATTTCTGG	21599
Qy	301	aacaggcaagcaatcacaaaaagctgltttacaacacataaggcgatlttactttcac	360
Db	21600	AACAGCGACAGCAATCACAAAAGCTGTTTAAACAACATAAGGGGATTTGACTTTCAC	21659
Qy	361	aggtacaggggaactctctatgtttccaaagctgtagtcgaagactgtctacaggggctgc	420
Db	21660	AGGTACAGGGAACTCTTATTGTTTCCAAAGGATGAGGAGCTGTAGCGAGGGGCTGC	21719
Qy	421	tgtaacagcagcgtgctgagataaactcacccagttataaggtttctctgcatacttt	480
Db	21720	TGTTACAGCAGCGGTGATGATTAATCTAACACAGTTTAAAGGTTTCTCTCCCTATCTT	21779
Qy	481	tattgcgtctctctggaagttcgataactacgcggcaaaagagccgttaagctctacggg	540
Db	21780	TATTCCGCTCTCTGGAAAGTTCATACTACACCAGCAAGAGAGCGCTTAACTGCTCTACGG	21839
Qy	541	tagcttggagtttgcacaaaaatgtaagtttgctcttcagcaaaaacttttcaacgagata	600
Db	21840	TAGCTTGAGTTTGAC-AAAAAGTCAAGTTGCTCTTTCACACAAAACCTTTTCAACGGATTA	21898
Qy	601	tggcggctctacacacgcaaaaactcttcattaaacagggactacaatgtcacgctcgt	660
Db	21899	TGGCGGCTCTATCACCGCAAAAACCTTTTCATTACAGGAGACTACAAATCTCAGCTGT	21958
Qy	661	ttctgaataatccctctcaaaagaaggcgagccattgaagcttcgatggccctaacat	720
Db	21959	TTCTAAAAATACCTCTCTCAAAABAAGGGGAGCCATCTCAAGCTTCGAGCCCTTACAT	22018
Qy	721	tacttgaacaccaaaggagagctctcttttcttacaatacttcttgcgattctctgaagctgc	780
Db	22019	TACTGGAACCAAGGGGAGAGTCTCTTTTCTGACAATACTTCTGGAGTCTGGAGCTGC	22078
Qy	781	aattttacagaagcctcggtgacattctctaataatgtcaaatgcttcccttattgacaa	840
Db	22079	AATTTTTCAGAGAGCCTGGGTACTATTCTTATATATGTTAAAGTTTCTTTATTGACAA	22138
Qy	841	taagttcacagagagagctctctcaacaacagggagatagttcaagaggtgacatctgac	900
Db	22139	TAAAGTCAAGAGGAGGACTCTCTCAACAAGGGGAGATATGTAGAGAGTGTCTATCTGTGC	22198
Qy	901	ttataaaactagtagacagataactaaagtlcaaccttcaacttggaaatcagatlttacttcaag	960
Db	22199	TTATTAAACTAGTACAGATACTTAAGGTCAACCTCACTGGAAATCGAGTGTACTCTTCAG	22258
Qy	961	caacaatatctcgaaaaaagcggggggactatctatgtgaaaaagcttcaactgcttc	1020
Db	22259	CACAAATCATGTGACAAACGCGGGAGAGCTATCTATGTGAAAAGCTTCGAACCTGGCTTC	22318
Qy	1021	cggagagacttaccctattctagtagaataatgttcaatggaagtacagctccctaaagctg	1080
Db	22319	CGAGAGACTTACCTTATCTAGTAAATATGTGTCAATGAGAGGTACAGCTCCATAAAGTGG	22378
Qy	1081	agccataactatcgaaagatagctggggaattgagtttataccgcgagatggtgtgacattgt	1140

D	22379	AGGATATGCTATGGAAGATATAGTGGGGAATTAGTATTATCCCGCATATGCTGATCATGT	22435
Q	1141	cttttagaggaaatacagctacttctaactcttcctggagcaatagaagtatgatacgactt	1200
D	22439	CTTTTATGGGAATACAGTACTTCTACTACTCTCTGGAGCAATATGAAGTATGTTCGACTT	22498
Q	1401	aggaaacagatgcaaaataatacagcttctgcgtctctctctgtagagcaatactactcta	1260
D	22499	AGGACAGATGCGAAATATACAGCTTTGCGTTCTGCTCTGTAAGCCATCTTACTTCTA	22555
Q	1261	tgatccataactacacagatccacacaacagttacagatgcttctaaagttaatgagac	1320
D	22559	TGATCCCATTAACATACAGATCATCTCCAAACAGTTTACAGATGTCTTAAAGTATATGAGAC	22618
Q	1321	tcggcgacatctgcgcataataatacagggaaacactctctcccaagagaaaaagttatc	1380
D	22619	TCCGGCAGATTCGTGACTCCAAATATACGGGAACATCATCTTCCACAGAGAAAAAGTTATC	22676
Q	1381	agaaacacagccgcagatctctaaaaacttctaactcgaaagctactacagcctgtaactc	1440
D	22679	AGAAACAGAGGCGCGAGATTCTTAAAAATCTTACTTCGAAGCTACTACAGCTGTAACTCT	22736
Q	1441	ttcaagagtgactctatctttaaacaatggagtgactctgcagactcaagcatctactca	1500
D	22739	TTCAAGGAGGTACTCTTCTTTAAAAATGAGATGTACTCTGACATCAGGCATCTACATCA	22798
Q	1501	acagcgacatcttcgctcctcgaatatgacgttagaacctacttagaacctctgatactag	1560
D	22799	ACAGCGCAGATTCGCTCGCTCGAAATGAGAGCTAGAAACTCTTAAGAACTGCTGTACTAG	22855
Q	1561	caccataaacaatttggtcattaaatcagatctataagacggtgcgaagaagccaaaat	1620
D	22859	CACCATTAAACAATTTGTCATTAAATCATGATCTCTATACCGGTSCAAAGAAAGCAAAAT	22918
Q	1621	agaaaccaaaagctacgtcaaaaaactctactttctcgaaacactcaacttattggacc	1680
D	22919	AGAAACCAAAAGCTACTCCTCAAAAATCTGACTTTATCTGAAACCAATCACTTATTGGACC	22977
Q	1681	gaaggggacagttttaaagaaatcactagtttaagaatatcctaagcctcctaagacatctaga	1740
D	22979	GAGGGGACACTTTTATGAATAATCATAGTTTAAAGAAATCCTCAGTCTTAAGCATCTTAGA	23036
Q	1741	gtccaaagctcttcgagactgtaacaagaacgcgagtgactccaaatccataatgggtga	1800
D	23039	GCTCAAAAGCTCTCGGAACGTATACAGAACCGCATGTACTCCAAATCTTAAATGGGTGA	23098
Q	1801	gaattccattacagcgtacatacagaagaacttggggccaaattcttcgggggacagggcttc	1860
D	23099	GAAATTCATTAACGGGTATACAGGAACCTTGGGGCCCAATTGTTGGGGGACAGGGCGTTT	23158
Q	1861	tacgaactgaacctcaactgaactgaataaaacttgcataatctcctaacccagacgatacgg	1920
D	23159	TACGACATGCACACTTCAATGTGACATAAACATGCGCTATTTCTTAATCCGACGGTATCG	23218
Q	1921	ctctcttagtccccaatagattatgaaatgacattatagatatcttgccctccataatct	1980
D	23219	CTCTTTTAGTCCCAATAGCTTTATGAAATGCACTTTATTAATATTAAGCTCTCTCCATTAATC	23278
Q	1981	tatgagaaactgcacaaagaaagggttcgaaggagaccggtcttcttggtgtctgtagtatac	2040
D	23279	TATGGAAGACTGCAAAAGAAAGGTTTCAGAGGAGACCGTCTTTGGTGTCTGGATTATC	23338
Q	2041	taactcttccataaagataagtaacaacaaacagacgcgggttccgcaattgtagtggcgg	2100
D	23339	TAACTCTTCCATAAAGATATGATACAAAAACAGACGGCGGTTTGCATTGTGATGGCG	23398
Q	2101	ttatgtcataaggaggaacctacataactgtcttcagataaagaattctagtcgcatcttg	2160
D	23399	TTATGTCAATAGAGGAACCTTCAATACTGTTTTCAGATTAAGATTCTTAGTGTCTGATTTTG	
Q	2161	tcagctctcttggaagagataagagactactcttgtagctaaagaatcaagtaacgtctg	
D	23459	TCAGCTCTTTGGAAAGATATAGATCTACTTTGTATGCTTAAGATCAAGGTACAG	

FEATURES	source
RESULT	5
LOCUS	AB1835
DEFINITION	AB1835
ACCESSION	Sequence 9 from Patent WO9858953.
VERSION	AB1835
KEYWORDS	AB1835.1 GI:6731868
SOURCE	unidentified.
ORGANISM	unidentified
REFERENCE	unclassified.
AUTHORS	1 (bases 1 to 2787)
TITLE	Madsen, A. and Birkelund, S.
JOURNAL	NOVEL SURFACE EXPOSED PROTEINS FROM CHLAMYDIA PNEUMONIAE
	Patent: WO 9858953-A 30-Dec-1998;
	MADSEN ANNA SOFIE (DK); BIRKELUND SVEND (DK)
	Location/Qualifiers
	1..2787

Qy	101	agaagctctcttccccaagtttgatatttctcaattgcbatttt	
Db	1	ATGAAGCTCTCTTCCCAAGTTGTATTTTCTACATTGCTATTTTCTCTGTATG	60
Qy	161	attaccacgagcaagtttggattcaagtggatttcgatggaaataaaagcgaat	220
Db	61	ATTGCTACGAGACAGACTTTTGGATTCMACTGGCAATTTCGATGGGAATAAAAAGGTAAT	120
Qy	221	ttttcaattcgtagagatcagaagaatgctgtgaactacctaactttaaagaaatgct	280
Db	121	TTTTCAATTGCTAGAGTACAGAAATGCTGGAATCCTACCTATTATTAAGGAAATGTC	180
Qy	281	acctgagaataatcttccttgaaacagcagcaagaatcacaaaaagctgtttaacaact	340
Db	181	ACCTGAGAAAATTTCTCTGAAACAGCACAGCAATCAAAAAACCTGTTTAAACAACT	240
Qy	341	aagggcgaattgactcttcacagtgaaacggaaactctcattgtctccaaacggtgtagca	400
Db	241	AAAGGCGAATTTGACTTTCACAGTAAAGGGAACCTGTATGTTCCAAAGGCGATGCA	300
Qy	401	ggagcgtgtgacgggggtctgtttaacagcagcgtygtlaataaactaccagttta	460
Db	301	GGAGCTGTAGCAGGGGCTGCTGTAAACAGCAGCGTGTGATTAATCTACCAAGTTTAT	360
Qy	461	gggtttcttcgctaacttttaatgtgcgtctcttgaaagttcgaataactaccgcaaga	520
Db	361	GGGTTTCTTCCGCTATCTTTATTTATTCGCTCTCTGAAATTCGATACCTGGCAAGA	420
Qy	521	gccgttagctgctctacggtagcttgagtttgacaaaatgcaagttgtccttcagc	580
Db	421	GCGGTAGCTGCTACGGGTAGCTTGAAAGTTTGACAAAATGTCAAGTTTGCCTTCAGC	480
Qy	581	aaaaattctcaaggataaaggcgggtgctatcacgcgcaaaaactcttcataacagg	640
Db	481	AAAACTTTTCAACGGATTAAGGCGGTCTATACCCCAAAAACCTTTTCATTAAACAGG	540
Qy	641	actacaatgtcagcgtctgtttcttgaaataacccctccaaagaaggcggaagcattcg	700
Db	541	ACTACAAGTACGCTGCTGTTTTCGAAAATTAACCTCCCAAGAAAGGCGAGCATTCAG	600
Qy	701	acttcgagtgcccttacattacttgaaaccaaagggaagtcctcttttctgacaattact	760
Db	601	ACTTCCGATGCCCTTACCTTAACGTGAAGAACCAAGGGGAAGTCTCTTTCTGCACATACT	660
Qy	761	tcttcggaattcggagctgcgaatttttaaggaaagctcggtgctatttctaataatgct	820
Db	661	TCTTGGGATTCTGAGCTGCATTTTACGAAAGCCCTGGTGCTATTTCTAATATGCT	720
Qy	821	aaagtttccttaattgacaataagttcacagagcgagctccctcaacaacgggggatatg	880
Db	721	AAAGTTTCCTTATTATGACAATAAGGTACAGAGCGAGCTCCCAACAACGGGGGATATG	780
Qy	881	tcaggaagtgctatctgtgttatataaactagtacagataactaaggtcaacctcaatg	940
Db	781	TCAGGAGGTGCTATCTGTGTTTAAAACTAGTACGATCTAAAGTCACTCACTGGA	840
Qy	941	aatcagagtgtaaccttcacacaacaataactcgacaacaacggggggggtctatctatg	1000
Db	841	AATCAGATGTACTCTTCACACAACAATACATGACACAACGCGGGAGACTATATGATGTG	900
Qy	1001	aaaaagctcgaactggtcttcggaggaactacccttaactagtagaataagtgtaatga	1060
Db	901	AAAAAGCTCGAATGCTGCTCCGAGAGCACTTACCTTTCAGTAGAATAATAGTCAATGGA	960
Qy	1061	ggtcagctcctaagaagtgagccatagctatcogaagaatgtygggaattgagttaacc	1120

Db 961 GGTACAGCTCCTAAAGGGAGCCATACCTATCGAAGATAGGGGAAATTAGATTATCC 1020  
 QY gccatagtggtgacatgctcttttaagagatacagctacttactactccctggagc 1180  
 Db 1021 GCCATAGTGGTGCATTTGTTTAAAGGAAATACAGTCACTTCTACTACTCTGGGAGC 1080  
 QY 1181 aatagaagtagtalcagacttagaagagtgcaaatagatgacgtcttgctgctgct 1240  
 Db 1081 AATAGAAGTACTATCGACTTAGAAGAGTCAAGATGACAGCTTGCTGCTGCTGCT 1140  
 QY 1241 ggtatagagcatcactactcttaagatcccaatacctcagatcatcccaacagttacagat 1300  
 Db 1141 GGTAGAGCATCTACTCTTATGATCCATTAATACAGATCTTCAACACAGTTACAGAT 1200  
 QY 1301 gctctaaagtaataagagaccccgcaaatctcagctacataataagaggaatac 1360  
 Db 1201 GCTTTAAAGTTAATGAGACCTCCGCAATCTGCACTACAAATTTACAGGAAATATC 1260  
 QY 1361 ttcaagagagaaagtatcagagacagagccgcaagatctctaaatcttacttcgaag 1420  
 Db 1261 TTCACAGAGAAAGTTATCAGAGACAGAGGCCGAGATCTTAAAAATCTACTTCGAG 1320  
 QY 1421 ctactaagcctgtaactcttccagaggtactctatctctttaaactgggtgactctg 1480  
 Db 1321 CTACTACGCTGTACTCTTTCAGAGGTACTCTATCTTAAACATGAGTGAATCTG 1380  
 QY 1481 cagaactcagcatcactcaacagagcagatctcgtctcgaatcagagcttgaagact 1540  
 Db 1381 CAGACTCGGCACTTCACTCAACAGAGAGATTCGTCTCGAAATGAGAGTGAAGTACT 1440  
 QY 1541 ctgaactcgtgtactactagacacataaacttggtcatcttaacatcagttctatgac 1600  
 Db 1441 CTAAACCTGCTGATACAGACCATTAACATTTGGTCATTAACATCAGTTCTATGAC 1500  
 QY 1601 ggtcacaagagagcaaaatagaaaccaaagctacgtctaaataatcgaacttctgga 1660  
 Db 1501 GGTCAAAAGAGGCAAAAATGAAACCAAACTACGTCAAAATAATCTGACTTTATCTGGA 1560  
 QY 1661 accatcacttatctgagcccgacagcggtcttatagaatacagtttaagaatcct 1720  
 Db 1561 ACCATCACTTATTTATGAGACCCGAGCGGCGTTTATGAAATCAATGTTAAGAAATCCT 1620  
 QY 1721 cagcctcactagacatcttagagctcaaaagctctggaactgtcaacaagacccgagtgact 1780  
 Db 1621 CAGTCCTACGACATCTTAGAGCTCAAAAGCTTCTGAACTGTACAAAGACCCGAGTGA 1680  
 QY 1781 ccagatccataatcgggtgagaaattccattacggctatcagggaaacttggggccaatt 1840  
 Db 1681 CCAATCTCTAATATGGGTGAGAAATTCATTCAGGCTATCCAGGAATCTTGAGGCCAAT 1740  
 QY 1841 gtttggggagagaggggtctcagactgcaactcctcaactgtaataaactgctatat 1900  
 Db 1741 GTTTGGGGAGAGAGGGGCTTTCACACTGCACTCAACTGACTTAAAGTGGCTATAT 1800  
 QY 1901 ccctatcccgagcgatccgctcttagctccctaatagcttatggaatgacttatagat 1960  
 Db 1801 CCTATCCCGAGCGTATCGGCTTTTATGCTTCCCTAATAGCTTATGAAATGCAATTAAT 1860  
 QY 1961 atagctctctcattatctatagagactgcaaacagaggtctgacagagagacgtgct 2020  
 Db 1861 ATTAGCTCTCTCATTAATCTATGAGAGCTCAAAACGAAAGGTTCAGAGGAGACCGT 1920  
 QY 2021 ttgtgggtgtgtatctatctaaactctcctcaatagatagatgacaaacaacagcgagg 2080  
 Db 1921 TTTTGGGTGTGTGATTAATCAACTTCTTCCATAAGATATACAAAACACGAGCGGG 1980  
 QY 2081 ttgcacatttgagtgagggtgtatgctatagagagaaacacacactactgttgcataag 2140  
 Db 1981 TTTCGCCATTGAGTGGCGGTATATGCTATAGAGAGAAACCTTCACTTGTTCAGATAG 2040  
 QY 2141 attcttagtgcgacttctgcaagcctcttggaaagatagagactctttagtctaaag 2200

Db 2041 ATTCTTAGTGTGCTGATTTTGTCTCAGCTCTTGAAGAGATAGAGACTTGTAGCTAAG 2100  
 QY 2201 aatcaagatcagctcctcagaggaactctctattaccaagacaagaaactatctc 2260  
 Db 2101 AATCAAGTCAAGTCTACGAGAACTCTCTATTCACGACAAACGAAACCTTATCTCT 2160  
 QY 2261 ctctctcgaactcagagcctctgctgtctctatgttctcctcagaagatctcgtctc 2320  
 Db 2161 CTTCCTGCAACAGAGCGCTTGTGTGTCTTATGCTTCTTACAGAGATTCCTGTTCTC 2220  
 QY 2321 ttctcaggaaccttagctcaccaccatacagatacagatctgaaacaaactatcaaca 2380  
 Db 2221 TTTTCAGGAACCTTACTACACCCATACCGAATACGATCTGAAACCAAGTATACACA 2280  
 QY 2381 tatctcactgttaaaggaagctgggggaatgatagttctgctttagatctggatggaga 2440  
 Db 2281 TATCCTCTTAAAGGAACCTGGGGAATGATGTTTCCCTTAGATTCGGTGGAGA 2340  
 QY 2441 gctccgattgcttgaatgaatgagtgctctatctgagcagatacagctcctcaatgaatg 2500  
 Db 2341 GTCGCCATTGTGTAGATGAAGAGTCTATTTGAGCAGATACATGCCCTCATGAATTCG 2400  
 QY 2501 cagttgtctatgacactcaggaaggttttaagaagaaggaacagagactcgtgaatt 2560  
 Db 2401 CAGTTGTCTATGCACATCAAGAGGTTTAAAGAAACAGGAAACAGAGCTCGTGAATTT 2460  
 QY 2561 ggaagtagcgcgtcttgaatcttgcctcctacatcagggatctgataaggaatca 2620  
 Db 2461 GGAAGTAGCGGCTGTGTGATCTTGCCTTACCTATCGGATTCGATTTGATGAAGATCA 2520  
 QY 2621 gaetgccaagatgcaacgtacaactcactctgtgtatacgtgtgatactgttcgtagt 2680  
 Db 2521 GACTGCCAAGATGCAAGTCAATCTTAACCTGTGTATCTGTGATCTGTGATCTGTGAGT 2580  
 QY 2681 aaccggactgttcgagaacacagcgaatagcgtgattcttggaaaacttcgagacg 2740  
 Db 2581 AACCCGACGTGAGAGAACACTGCGAATTAAGCGGTGATTTGGAAAACTTCGGTACG 2640  
 QY 2741 aattggcaagacaagctttagctcctcgtcgtcagggaaacatttgccttaactcaat 2800  
 Db 2641 AATTGGCAAGACAAGTTTAGTCTTCGTCGAGGAGAACATTTTCTTAACTCAAT 2700  
 QY 2801 ttggaagccttggccaattctcttggaaatgcgtggtgcatctcgcgaattcaatgta 2860  
 Db 2701 TTGGAACCTTTGGCCAAATTTCTTTGAATGCGGTGATCTGCAATTAATGTA 2760  
 QY 2861 gacttagagcaaatatacactctca 2887  
 Db 2761 GACTTAGAGCAAAATACCAATTTCTAA 2787

RESULT 6  
 AB033795  
 LOCUS 2241 bp DNA BCT 31-OCT-1999  
 DEFINITION Chlamydomophila pneumoniae pmp-3.2 gene for polymorphic outer  
 membrane protein G family, complete cds.  
 ACCESSION AB033795  
 VERSION 1 GI:6172316  
 KEYWORDS polymorphic outer membrane protein G family.  
 SOURCE Chlamydomophila pneumoniae (strain:J138) DNA.  
 ORGANISM Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.  
 REFERENCE 1 (sites)  
 AUTHORS Shirai,M., Hirakawa,H., Ouchi,K., Tabuchi,M., Kishi,F., Kimoto,M.,  
 Takeuchi,A., Nishida,J., Shibata,K., Fujinaga,R., Yoneda,H.,  
 Matsushima,H., Tanaka,C., Furukawa,S., Miura,K., Nakazawa,A.,  
 Ishii,K., Shiba,T., Hattori,M. and Kubara,S.  
 TITLE Comparison of outer membrane protein genes omp and pmp in the whole  
 genome sequences of Chlamydia pneumoniae isolates from Japan and US  
 JOURNAL Unpublished (1999)  
 REFERENCE 2 (bases 1 to 2241)  
 AUTHORS Hirakawa,H. and Shirai,M.  
 TITLE Direct Submission

## JOURNAL

Submitted (25-OCT-1999) to the DDBJ/EMBL/GenBank databases. Hideki Hirakawa, Kyushu University, Graduate School of Genetic Resources Technology; 6-10-1 Hakozaki, Higashi-ku, Fukuoka, Fukuoka 812-8581, Japan (E-mail: hirakawa@gtl.kyushu-u.ac.jp, URL: http://www.gtl.kyushu-u.ac.jp, Fax: 81-92-642-3043).

## FEATURES

Location/Qualifiers

1..2241  
 /organism="Chlamydia pneumoniae"  
 /strain="J138"  
 /db\_xref="taxon:83558"  
 /note="synonym: Chlamydia pneumoniae"  
 1..2241  
 /gene="pmp-3.2"  
 1..2241  
 /gene="pmp-3.2"  
 /note="pmp-3.2"  
 /note="containing frame shift"  
 /codon\_start=1  
 /transl\_table=11  
 /product="polymorphic outer membrane protein G family"  
 /protein\_id="BA85948.1"  
 /db\_xref="GI:6172317"  
 /translation="MSALFSENTSKKGAIGTSDALITNGOEVSESDNTSPSGA  
 ALPTASVTSINNAKVFIDNKVTGASSTTGMSGGAICATKSTIDTAVTLTGOML  
 LFSNMTSTAGAIYKLELAGGLTFSRNSVNGTAPKGAIAEDSGELISAD  
 SGDIVLGNTVSTPTGTRSSIDISTSAKMTAKRAARAIYFTDPIITGSSSTVTD  
 VLKNETPADSAIQGNITFTGEKLSIEADSKNLSKLPVLTLSGSLSLHGV  
 TQTOCAFTQOADSRLMDVGTLEPADISTINLVINISIDGAKAKLETAKTNL  
 TSGITTLDPTEFTFENSLRNPOSDYLIELKASGVTSIATVPDINGEFHYGO  
 GWGPIWGTGASTATAPMTKGTGIPNPERIGSLVPLNMAFIDISLHMETAN  
 EGIQDRAFWCAGLSNFPHKDSIKTRGRFRLSGGVIGNLTGSDKLSAIPOLF  
 GRDQYFAVNAKOGTGVGLYIOHNETYISLPCKRPSLSVPEPTEIVLSEGNLSIT  
 HDNDLKTYYTTPYTKSGMGNDSFLEFGSAPICLDSALFEDIMPMKIQFYIAH  
 OSGFKGTEAREPFGSSRYNVALPIGIFRDESDQDNTVLTGIVDVLRSNPD  
 TTTLRISGDSMKTFTGNLAKRALVLRAGNHFCFNSFNAFQSFEELGSSRHYNDL  
 GAKIQP"

## CDS

BASE COUNT 660 a 474 c 478 g 629 t  
 ORIGIN

Query Match 74.6%; Score 2239.4; DB 1; Length 2241;  
 Best local Similarity 100.0%; Pred. No. 0;  
 Matches 2240; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 647 atgtcagctctgtttctgaataacccctcaagaaggcgagccattcaagcttc 706  
 Db 1 ATGTCAAGCTCTGTTTCTGAATAACCTCTCAAGAAAGCGAGCCATTCAAGACTTC 60  
 QY 707 gatgcccctaccctactctggaacccaagggaagtcctctttctgacaatactcttcg 766  
 Db 61 GATGCCCTTACCAATTACTGGAAACCAAGGGAAGTCTCTTTTTCGACAAATCTTCTCG 120  
 QY 767 gatctcgagctgcaatttttcaagaagcctcggygacatttctaatactgaagtc 826  
 Db 121 GATTCGGAGCGTCATTTTACAGAGAGCTCGGTGACTATTCTTAATAATCTTAAAGTT 180  
 QY 827 tcccttattgacaataagtcacagagcgagcctctcacaacaggggatattgacaga 886  
 Db 181 TCCTTATTGACAATAAGSTCACAGAGGAGACTCTCTACACAGCGGGGATATTGACAGA 240  
 QY 887 ggtgcatatggtcttataaaactgtacagataactaagtaaccctactggaatacag 946  
 Db 241 GGTGCTATGTGCTTATAAACTAGTACAGATTAAGTACCTCACTGAGAAATCAG 300  
 QY 947 atgttactcttaagaacaacatcacgacaacagcgagagagagctattatgtgaaag 1006  
 Db 301 ATGTACTCTTACAGACATATCATCGACACAGCGGGAGAGCTATATGTGAAAG 360  
 QY 1007 ctgcgaactggctccggagagactaccctattcagtagaataagtagtcaatggaagtagca 1066  
 Db 361 CTCGAACTGGCTTCGGAGAGACTTACCTATTCACTAGAAATAGTGCATAGAGAGTCA 420  
 QY 1067 gctcctaagtgtagagcatagctatcgagaagatagtggaattcgttataccgcagat 1126

Db 421 GCCTCTAAAGTGTGAGACCCATAGCTATCGAAGATAGGGGAATTAGTTATCCGCCGAT 480  
 QY 1127 agtggtagaattgctcttttaaggaataagtaagtaagtaagtaagtaagtaagtaagta 1186  
 Db 481 AGTGGTAGAATTGCTCTTTTAAAGGAATACAGTACACTTCTACTACCTCGGAGCAATAGA 540  
 QY 1187 agtagtagacttaggaacgaagtagcaagaatgagaagcttgcgtctgcgtgtaga 1246  
 Db 541 AGTAGATAGACTTAGGAAGAGTGCAGAAAGTAGAAGAGTTCGCTTCGCTCGTGTGAGA 600  
 QY 1247 gccatctacttctatgatccataactcaagatcatccacaacagttacagatgctcta 1306  
 Db 601 GCCATCTACTTCTATGATCCATTAAGTACAGATATATCACAACAGTTACAGATGCTCTTA 660  
 QY 1307 aaagttaatgaaactccggcagagattctgaactacaataatacagggaacatcatctaca 1366  
 Db 661 AAAGTTAATGAGACTCCGGCAGATTTCTCAGTAATATACAGGGAACATCATCTTACCA 720  
 QY 1367 ggagaaagttatcagagacagcgagagctctctaaatcttacttgcgaactacta 1426  
 Db 721 GGAGAAAGTTATCAGACACAGAGGCCGCAAGATCTTAAATCTTACTCGAACTCTA 780  
 QY 1427 cagcctgtaactcttcagagagttactctatctttaaacaatggaagtagtgcagact 1486  
 Db 781 CAGCCTGTAACTCTTTCAGAGAGTACTCTATCTTAAACATGGAAGTACCTCTCAGACT 840  
 QY 1487 cagagattcatcacaacagcgagattctctgctgaatgagcgtgaagaactactctagaa 1546  
 Db 841 CAGGATTCACCTCAACAGAGAGATTCCTGCTCAAAATGACGTAGAGAACACTCTAGAA 900  
 QY 1547 cctgtctatcagacacataaacaatttggcataacacagttctatagaagagtgca 1606  
 Db 901 CCTGTGATAGTACAGACATTAACAATTGCTATTAACATCACTTTATAGAGAGGTGCA 960  
 QY 1607 aagaagggcaaaaatagaacaacaaagctacgtcaaaaatctgacttattctggaaccatc 1666  
 Db 961 AAGAAGGCCAAAATAGAAACCAAGCTACGTCAAAAATCTGACTTATCTGGAACATC 1020  
 QY 1667 acttattgagcccgagcgagcgtttatgaaatcatatggttaagaatctctcagtc 1726  
 Db 1021 ACTTATTGAGCCGAGAGGCGACCTTTTATGAAATATATGTTAAATAATCTCTAGTCC 1080  
 QY 1727 taagacatcttaagagctcaagagctctggaactgtaacaaacccgagtagtccagat 1786  
 Db 1081 TACGACATCTTAGAGCTCAAAAGCTTCTGGAAGTATACAGACACCGAGTACCTCCAGAT 1140  
 QY 1787 cctataatgggtggaatccatcttaagcgtatcagaggaacttggggcccaattgtttg 1846  
 Db 1141 CCTATAATGGGTGGAATTCATTAAGGCTATCAGGGAACCTTGGGGCCCAATTGTTGG 1200  
 QY 1847 ggagcaaggagcttcaagatgcaaccttcaactggaactaactggcctatatctctaat 1906  
 Db 1201 GGAGACAGGGGCTTCTAGATGCACTTCACTGACATAAAGCTGCTATATCTCTAAT 1260  
 QY 1907 ccgagagctatcgagctcttaagtcctcaatagctttagaatgcatattatagatattagc 1966  
 Db 1261 CCCGAGGCTATCGGCTCTTATGCTCCATAATAGCTTATGGAATCATTTATAGATATTAGC 1320  
 QY 1967 tcttcattatcttctgagactgcaacacgaagggttgcagggagacgctgcttttgg 2026  
 Db 1321 TCTCTCATTAATCTTATGAGACTGCAAAAGAGGGTTCAGGGAACCGTGTCTTTTGG 1380  
 QY 2027 tgtgtgattatctactctctcacaataagtagaacaacaaacagacgaggtttgc 2086  
 Db 1381 TGTGCTGATTATCTAATCTCTCTCCATAAGATAGTACAAAAACAGACCGGGTTTCC 1440  
 QY 2087 catttagtgaggttatatgcataggaagaaacacacactattgttcagataagaattctt 2146  
 Db 1441 CATTTGATGTCGGGTTATGATCATTAGAGGAACCTACATACCTTGTTCAGATTAAGATTCT 1500  
 QY 2147 agtgcagcatttgtcagcctcttggaaagataagagactattgtgtcagaatacaaa 2206

Db	1501	AGTGGTCAGTATTGTGCAGCTCTTTGGAAAGATPAGAGACTAATGTTGTACTAAGAAATCA	1560
Qy	2207	ggtaacagcttaacggaggaaactctctattaccagcacacgaaactatactctctct	2266
Db	1561	GGTAAAGTCTACGGAGGAACCTCTCATTTACCGACACAGAAACCTATATCTCTTCT	1620
Qy	2267	tgcacaactacggccctgttcctgtctcattatgtcttcacagagatctctctcttca	2326
Db	1621	TGCAAACTACGGCCCTTGTCTGTCTTATGTCTTCTTACAGAGATTCCTGTCTCTTTTCA	1680
Qy	2327	ggaaacctagcttacaccataccgataacgactctgnaaaccaagtatatacaacatacct	2386
Db	1681	GGAAACCTTAGCTACACCCATCGGATACGATCGATCTGAAACACCAAGTATCAACATATCCT	1740
Qy	2387	actgttaaggaagcttgggggaatgatagtttcgccttttgaaattcggtgtgaagactccg	2446
Db	1741	ACTGTTAAAGGAAGCTGGGGGAATGATATATTCGCTTTGAATTCGGTGGAAAGCTCCG	1800
Qy	2447	attgcttagatgaaagtgctcattattgagcagiatcatgcccctcatgaattgcagttt	2506
Db	1801	ATTGCTTAGATGAAAGTGTCTCTATTTGAGCAGTACATGCCCTTCAATGAATTGCAAGTTT	1860
Qy	2507	gtctatgcacatcaggaaggttttaagaacaggaaacgaagcttcgtgaatttggaaat	2566
Db	1861	GCTATGCAACATCAGGAAGGTTTAAAGAACAGGGAAACGAAAGCTGTGAATTTGGAAAT	1920
Qy	2567	agccgtctgttaactcttccctaccataccgagatccgatttgcataaggaatcacagctc	2626
Db	1921	AGCGTCTTGTGAATCTTGTGCTTACTACTACGGGATCCGATTTATPAGGAATCAGACTGC	1980
Qy	2627	caagatgcacagctacatctaacctcttggttatactgtgagctctgttcgtgaacccc	2686
Db	1981	CAAGATGCAACCTACATCTAATCTATCTGTGTATACGTGAGACTTGTGTGATTAACCCC	2040
Qy	2687	gactctacacacacactggaattgacggttgattctcttgaaaaacttcgtagaatgtg	2746
Db	2041	GACTGTACGACACACTGCGAATTTAGCGGTGATTTCTTGGAAACCTTTCGTGCAATTTTG	2100
Qy	2747	gcaagacaagctttagcttccttcgtcagaggaaacaaatttctgtcttcaactcaaatltgaa	2806
Db	2101	GCAAACCAAGCTTTAGTCTCCTCGTCAGAGGAACCAATTTTGTCTTAACCAATTTTGAA	2160
Qy	2807	gcctttaacccaattctcttttgaattcgttggtgtaatctccgaaattacataatgaactta	2866
Db	2161	GCTTTAGCAATTTCTTTGAATTTGCTGGGTATCTCGCAATTTACATGTAGACTTAA	2220
Qy	2867	ggaagcaaaataccaattctaa 2887	
Db	2221	GGAGCAAAATATCAATTTCTAA 2241	
RESULT 7			
CPUT2499			
LOCUS	4926 bp	DNA	13-MAY-1999
DEFINITION	Chlamydomophila abortus putative outer membrane protein gene, partial cds; and putative 98 kDa outer membrane protein gene, complete cds		
ACCESSION	U72499		
VERSION	U72499.1	GI:1657776	
KEYWORDS	Chlamydomophila abortus.		
SOURCE	Chlamydomophila abortus.		
ORGANISM	Bacteria: Chlamydiales: Chlamydiales: Chlamydiales: Chlamydiales		
REFERENCE	Longbottom,D., Russell,M., Dunbar,S.M., Jones,G.E. and Herring,A.J.		
AUTHORS	1 (bases 1 to 4926)		
TITLE	98kDa protein genes from ovine abortion strain S26/3 Chlamydia psittaci		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 4926)		
AUTHORS	Longbottom,D.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-SEP-1996) Moredun Research Institute, 408 Gilmerton		
FEATURES	Road, Edinburgh EH17 7QH, Scotland, UK		
	Location/Qualifiers		

```

source
1. .4926
/organism="Chlamydomophila abortus"
/isolate="ovine abortion S26/3"
/species="host"Ovis aries"
/db_xref="taxon:83555"
/clone="S112"
/dev_stage="elementary body"
/notes="submitted as Chlamydia psittaci"
1. .1950
/codon_start=1
/transl_table=11
/product="putative outer membrane protein"
/protein_id="AAB187.1"
/db_xref="GI:165777"
/translation="DPITTPAPAAAVYTAQNPENSLAKATGSPAGSGAVSPAPAP
TKTPLEKIAPIPTQDEIQVAEAAQOQSAAYVNGKIVTSGEGLSEDDKKNPLNATSVTH
NDVSEIADQILVSGAGLIVDFQKQESLIVNGGTSIIINVTAPSEGLQSRSTPS
PKNAIPVIRAVASKVIASLLINLERADGSAVAVPTIEESDGSITITNLAVDNLSIE
NGKATPILAAKGGSGSVTLTGDLQFODSSONTYDNPDLNKNFTLNFDISAPADKIH
EGENIIIPGQATSNUNGCGKKEVNEVDSCKVSEPEKMYSTGIVIPANRRATIVPVS
VWCSAIDKRAQONLVEVTEEGEDFTRGIMIGISINFTFKDSTKVQEGRRITSSQVQ
VSTQISDKVNDLAFQMLGKSKDYRLADARSHVVAASIHAKCEKLVNHTFSSRKGA
ILARPEKSPITFPAQLSIVSLSHNSMTKTHPNPSSSRKNAHCVAGELGSLYLEILD
HPALIEPPEVKALIVVEOQDEPKETQGTGTRNFOSSAHFVNSILPGVGREKRNKN
TINILATQPDIDYRDAKSVKELPSVHTAMSTGATNLSRQMLIDGSDHHHTLDNLELV
FCHGAFELRGSSRNNDIGGRYKF"
2083. .2089
/notes="putative"
2094. .4874

```

	stem_loop	4884..4905	/note="potential 9-base terminator sequence"
BASE COUNT	1472 a	1033 c	1031 g 1390 t
ORIGIN			
Query Match		18.2%;	Score 546.6; DB 2;
Best Local Similarity		53.7%;	Pred. NO. 3.6e-107;
Matches 1390; Conservative		0;	Mismatches 1134; Indels 63; Gaps 10
QY	328	tttttaacaacttaagggcgatttgaacttcacaggttaacgggaactctctattgtcca	387
Db	2324	TTTTTCTTCAGATCGACAGAACCTTAACCTTCAAAGGGAACAACCAATAGCTTATCCATAAC	2383
QY	388	aacggttgatgtcaggagacgttagcagcgggctgctgttaacagcagcgtgtgtataatc	447
Db	2384	GAACGGAAATGCCGGAGCTAATCC---TCCGGGAATTAACTTAACACACGCCGATAGAT	2440
QY	448	taccagcgtttaaagggtttcttcgcctacctttattatgcctcctgtgaagttcgataac	507
Db	2441	TCTTAACGGCTGAACAGATTTTCTTAAGCTTGAAGCTTAAGGAATGCCCATCTTCTTAAGTAA	2500



QY	508	taccgycaaaggaagccgttagctgctcctaagcgttagcttgaattgacaaaatgctcag	567
Db	2501	TACTGGAAAAGGGGG---TATGAATCCGGAGGAGCATTAACCTATAGCCAAATATCCAG	2551
QY	568	ttgcgcttcagcaaaaacttttcaagcgaataagcggcgtcatcacccgcaaaaactt	627
Db	2558	TMTCTGTTTGAATCAGAACATTTCCGCTGAATATGGGAGCCATCTCTTGGCAACCTTT	2611
QY	628	ttcattacaaggagactacaatgtaagctctgttttctgaaaatactctccaagaag	687
Db	2618	TTCTCTAACCGGCTCGAGCAAAAGAAATCAGCTTACCACCTAACTCTACGCCAAAAAG	2677
QY	688	cggagccattcagcttcgcagtcgccttaccatctcggaaaccaaggaggagctctt	747
Db	2678	TGGACGATGCTGCTCAAGGAATAGCTATCTTCGGACAAACGAAGCAATCAGAT	2737
QY	748	ttctgacaatactcttcgatatctcggagctgcgaattttacagaagctcgtagcat	807
Db	2738	TTCTGGGAACACTGCTGTGAATCTGGGGGAGCGTATTCAGAAAGCTTCTATGACAT	2797
QY	808	ttctaatagcttaagttccttatttgacaataaggtcacagaagcgaagctctcaac	867
Db	2798	TTCAGGTAAACAACACGTTGCTTTTACCAACAATGCTGTTCCGGTTCATCT-----	2849
QY	866	aacgggggatagtcaggaaggtgctatctgtgcttaaaaacatgacagatactaagt	927
Db	2850	-----GATGGTGGCGGTGAGGACTATTCATGTGAGAAAACGTTGACACCACTT	2902
QY	928	caccctccttgaaatcagatgtaagttactcttcagcaacaatacgcacaacagcggag	987
Db	2903	TACTTAAAGATTAACAAGCTTTGATTTTGGAGAAATACTCTTCGCAAAAAGTGG	2962
QY	988	agctatactatgtgaaaaagctcgacagctgcttcggagagcttaccctatcagtagaa	1047
Db	2963	AGCGATTTACACCGATTAACATCATATGACTTTGTGGGGCCCTACGGCATTTATATAAA	3022
QY	1048	tagtgcatactggaagttcagctccttaaggttgaagccatagctacgaagatgtyggaa	1107
Db	3023	CAAAATTACCATCTCTCA---CCTAAGGTGGAGCTATTGGATATGCTGCCAAATGAGGA	3079
QY	1108	attgaattacccgcagatagtgtagactgtgcttttaaggaa---taagttactc	1164
Db	3080	ATGTAGCTTAACCGCTCAACATGAGGATATTACTTTGTAATAACCTGATGGCCACACA	3139
QY	1165	tactactcctcggagcgaatgaagtagtgcagcttagaagcagtgcaaatgacagc	1224
Db	3140	AGACAATGCTCAATTAATAAAGAAATGCAATTACATTGAAGCAATGTAATATGCTCA	3199
QY	1225	tttgcgttcgctgctgttagagccatctactctctctgtgctccataactacagatc	1284
Db	3200	CTTAGTGTCAGCGCTGGAAGAAGCAATTTCTTCTATGATCTTACAAAGTTG-----	3252
QY	1285	cacaaagttacagatgctcttaaaagttaagagactccggagatttgcatacata	1344
Db	3253	-----AAGGTAAATGCTGCTGATCTTCTCACTTTGATTAAGCTGAGGGTGAATAAGTA	3307
QY	1345	tacagggaaatcatctcttcacagagaanaagttacagagaagcgcgcagattctaa	1404
Db	3308	TAAATGAAGAATTAATTTTTCAGGAAAAAGCTCACTGAAGAACAAGTGTGTTGGGA	3367
QY	1405	aaatcttacttcgaagctactacagcctgttaactctttaggaagttactatctttaa	1464
Db	3368	TAACTTAAACACAACATTTACACAGCCTTACCTTAGTGCTGTGTAACCTTGTTTNCG	3427
QY	1465	acatgaggttagctctgcagactcaagcatctcaacagcagattctcgctcgaat	1524
Db	3428	CAGCGGTGTGAAGTGAAGCAAAAACAATCCTGTGCAACACAGAGATCTTTGATTCGAT	3487
QY	1525	ggacgttaggaactactcta---gaacctgtatctactagcaccataaacaattggtcat	1581
Db	3488	GGATGCGAGGCAAAAGTTATTCGCCAAAAACAAAGATGCTACACTGACGAATCTGGCTAT	3547
QY	1582	taacatagttctcatgaagcgtgcaagaagaagcaaaaatagaaccaagaagttatgtcaaa	1641

Db	3548	TAAATCCGAAATACCTTATGATGGGAAAAAATTTGGCCGTATGATGACCTTGTCTCTGGGAA	3607
QY	1642	aaatctgacttatactcgtgaaccatcaacttlatcttgaccacgagcgacgttttatgaaa	1700
Db	3608	GAATGTGACTTATTCAGGGGCTATTTGGGCTATTGATTCCTACAGGAAAGTTTATGAAA	3667
QY	1702	tcaatgttaagaabaatccacgtcctctaaagaactcttgagctcaagaactctcgaact	1767
Db	3668	CCATTAAGCTTAATGATACCTTAGCTTTAGAGGAATTCACATCTTCTGGGAAGGTTCGGT	3727
QY	1762	aaacaagaccgacagctgacccagatccataatcgtggtgagaabaatccatagcgcatca	1822
Db	3728	GACAAACAACCAAGTGGCCATGACATGTTGTGTGGTTCGCGAAACCCACTATGTTATTA	3788
QY	1822	gggaac-----ttggggcccaattgtttggggacaggggtctctacgactgc	1865
Db	3788	AGGAAACTGTGTGTGCAAGTTGGGTCAAAAGATTAATACCTGATCTCAAAAACAACAAACAC	3844
QY	1870	aaacttcaactgactaaacactcgtctatatccctcaaccccgagcgatcgctcttga	1922
Db	3848	AATCTTACTCTGGATTAATAACAGATATGTTCCAAATCCGAACGTCGTCCGCTGAT	3907
QY	1930	ccctaataagcttaatgaaatgcatltaataagatbaatagctctctccatlaacttataagag	1988
Db	3908	ACCAATATACCTTTGGGATCCTTTATAGATTATACGTTATCAAGATGTCTGGAAAGC	3967
QY	1990	tgcacaacgaaggttgc---aggagaacgctgtcttcttggctgagcttatcaact	2046
Db	3968	TAGTGTGATGATATCTTGTGAGACACGTCTGTGTTGTGGGTCTCTGGAAATGGGAACCT	4027
QY	2047	cttccaataagtagtacaaaacacgacgacgaggtttcgccatctttagagcggtatgt	2108
Db	4028	CTTCAATAAAGATCCGAATGCTGAAATCCCAATTCGCATATCCGATATACGTTGGGATATGT	4087
QY	2107	catagagagaacccatacactctgttctaataagaatctttagtgcgtcaactttgtcagc	2166
Db	4088	GTTAGAGCCACACAAATATACCTCGAAGAGGATCTCTTAGTGTGGCTTCTGTGCACTT	4147
QY	2167	ctttgaaagagatagagactactcttctagcctaagaatcaagatcagctctacgagagac	2226
Db	4148	ATTTCAAAAGTAAAGACTACCTGTATACACAGAAACGCCCAACGCTATACGGGGTTC	4207
QY	2227	tctctattaccagca-----caacgaacactatctctcttcccttgcaacatacggcc	2280
Db	4208	TCTATATTAATGAGATGTGACCAAGTTTGATGATCTCAACGCGTATTTAAAGGGCTTAA	4267
QY	2281	tgttcgctgttccttatgtctccacaagaagatctcgtcttccttttaagaabaacttagta	2340
Db	4268	CACGTGTGTTCAAGGTTTTCATAAAGATATCTCATATTTCTTGATGTCACAAATTAACCTA	4327
QY	2341	caaccatacgagataacgaatctgaaaccaaagtatacaacataccactagttaaagaag	2400
Db	4328	TTGGCACACGGCCAAACAACATGACACAGTCCTTTACAGACTATCTGAAGTGAAGAAAGTTC	4387
QY	2401	ctggggagataatagtttgcctttagaatctcggttgaagaagctcggattctctagaaga	2460
Db	4388	TTGGGGTAAATATACCTTGCGCTTTAACTTGTCTACTACCGTAACTATCCCGGATTTAG	4447
QY	2461	aagtgctctatcttgacgagatcacatgccccttcatagaatcagacgagtttgctctgacatca	2530
Db	4448	TTCTCTCTATCTTTGATGATGATATGACACCGTTTGCAAAATATACAAAGTGTCTATGGGCACA	4507
QY	2521	ggaaggttttaagaacaagaagacagaagctcgttgaatttggaaatagcgtcttgtga	2580
Db	4508	AGATGACTTTAAACAAACAACAACAAGCCGGGCTTTTGAAACACACGATCTTCTTCA	4567
QY	2581	tcttgcccttacctcgtgagctcgtattgtataagaatcagaatcgaagtcgaagtgtgaagta	2640
Db	4568	CGTTTCGTATACCTTAATAGTATAAATTTTGAAGAACTCTTCCTATGAGAGACAAAGTCTTA	4627
QY	2641	caatctaaactcttggtatatactcgttgatcttgcttgatgatacccgactttagacaaac	2700



Db 4628 TCATCTTACCTGATGATATATACCTGATGATACCGTCAATTAACCAAGCTGATGACAG 4687  
Qy 2701 actgcgaatgaacggtgattcttggaaacaccttgcgaatttggcaagaagaagctt 2760  
Db 4688 ATTGGGAGCATGACGACTTCTGCTGTTAACACACAGCTACGAACTTCTGACCAAGCTT 4747  
Qy 2761 agtccttcgtcgaaggaaacatttcttctaactcaatttgaagcctttagccaatt 2820  
Db 4748 CATAGTTCGCGGGGATACCATATGCTTACCTGCTGTTGAGATGTCAGACAGT 4807  
Qy 2821 tctcttggattgctggtgctcttcgcgaattacattgaaagccttgaagcaataacca 2880  
Db 4808 TCGTTTCGAATTAACGAAGCTTCAAGAAATTAACGATGATCTTGCGGCTAAGCTCC 4867  
Qy 2881 attctaa 2887  
Db 4868 GTTCTAA 4874

RESULT 8  
AE002192 12676 bp DNA BCT 30-MAY-2000  
LOCUS Chlamydomonas reinhardtii AR39, section 25 of 94 of the complete  
DEFINITION genome  
ACCESSION AE002192.1  
VERSION AE002192.2 GI:8163405  
KEYWORDS  
SOURCE  
ORGANISM Chlamydomonas reinhardtii AR39.  
REFERENCE Chlamydomonas reinhardtii AR39.  
AUTHORS Read, T.D., Brunham, R.C., Shen, C., Gill, S.R., Heidelberg, J.F., White, O., Hickey, E.K., Peterson, J., Umayam, L.A., Utterback, T., Berry, K., Bass, S., Linher, K., Weidman, J., Khouli, H., Craven, B., Bowman, C., Dodson, R., Winn, M., Nelson, W., Deboy, R., Kolonay, J., McClarty, G., Salzberg, S.L., Eisen, J. and Fraser, C.M.  
TITLE Genome sequences of Chlamydia trachomatis MOPn and Chlamydia pneumoniae AR39  
JOURNAL Nucleic Acids Res. 28 (6), 1397-1406 (2000)  
MEDLINE 2 (bases 1 to 12676)  
REFERENCE 2 (bases 1 to 12676)  
AUTHORS Read, T.D., Brunham, R.C., Shen, C., Gill, S.R., Heidelberg, J.F., White, O., Hickey, E.K., Peterson, J., Umayam, L.A., Utterback, T., Berry, K., Bass, S., Linher, K., Weidman, J., Khouli, H., Craven, B., Bowman, C., Dodson, R., Winn, M., Nelson, W., Deboy, R., Kolonay, J., McClarty, G., Salzberg, S.L., Eisen, J. and Fraser, C.M.  
TITLE Direct Submissio  
JOURNAL Submitted (01-MAR-2000) The Institute for Genomic Research, 9712  
COMMENT Medical Center Dr. Rockville, MD 20850, USA  
FEATURES  
source On Jun 1, 2000 this sequence version replaced gi:7189226.  
Location/Qualifiers  
1..12676  
/organism="Chlamydomonas reinhardtii AR39"  
/strain="AR39"  
/db\_xref="taxon:115711"  
/note="synonym: Chlamydia pneumoniae AR39"  
/gene="CP0300"  
complement(90..185)  
/gene="CP0300"  
/note="hypothetical protein; identified by Glimmer2;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AA38157.1"  
/db\_xref="GI:7189227"  
/translation="MFTSLGRSAIFFTGSLKHLRGLFELIFQK"  
complement(275..1819)  
/gene="CP0301"  
complement(275..1819)  
/gene="CP0301"  
/note="similar to GP:4376735; identified by sequence

gene  
CDS  
similarly; putative"  
/codon\_start=1  
/transl\_table=11  
/product="polymorphic membrane protein G family"  
/protein\_id="AA38158.1"  
/db\_xref="GI:7189228"  
/translation="MTLRNLFELCSALFLALPAAQVYVHESDGYNGALINRSLERK  
ITCYPECTSYIFLDDVRIISNVKHOEGAVFIRNSGMLPFMGRCNFTFNNLTGEG  
AATSRNGDTTLISNRSYALFASAPLPGCCALYLSGVMIENSEVYFCNYSM  
SGAATPYPLIGSKASRPVNLGNRLVIRDNVSGQGAISTHNTLTTRPSCSE  
NNHAYHVNNGGALAIAPGGSISIVSGDLFEKGFAPADQNTINSHLDGQAF  
KNHAYSEGVYEDPISHSEKRTIDLVINADEKETEYGTISFSGICLDHEVQAE  
NLSTIIDQVTLGAGTSLSDGYTLQHSFKQASSTLMSPTGLKCSGDVAONH  
ILIEDTNEFVPRIRAEKDALVLEKTLKAFEAQVMSYVDFPQEKATFPLEILG  
SPSLIGETTLERTQVTTENDAVRGWSLWSEYPPSLDKDRIRITPKTYVLTWNP  
EIRSNP"  
complement(2077..4926)  
/gene="CP0302"  
complement(2077..4926)  
/gene="CP0302"  
/note="similar to GP:4376733; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="polymorphic membrane protein G family"  
/protein\_id="AA38159.1"  
/db\_xref="GI:7189229"  
/translation="MYLRCFYIDANSLSKRSITMKTSIPWLVSVLAFSCHLSL  
ANBELSPDPSFNGNIDSGTFTPKTSATITSLGVDFFTEPGKGLPISDCFRQTDN  
LTFEGNHSITFEPTIDAGTAGAASTARKNTLFGSFLSDSSPSVTYTGQGL  
SSAGVNLNIRKLVANFSTADGAGKASFLTGTSDDALFNNSSSTKGAINT  
TAGARLNNYVRLFNINASTGALIDGDTGSLINNKRLTEGNAKYTGGAICMT  
KASGPELISNNKTLFASVNAVETSGAHHAKILSSGFEPLNNVNSAPKGG  
AISDAAGELSAENGTFFVNTLTGTSSTPKNRATNIGSGKTEPLRAKNT  
IFPYDPTSGTSGSDYIKINGSAGALNLPQGLTISGELTIDELKAVNLSVSTQ  
PVLISGKLLQGVLESTISFQESGLLGMOSGTLTSTASTITNTGIVNDSIG  
LKQPVSLTAGASAKRYVSGKLLIDEGNIVSHMSHQGLSLKTLTIDVADQNTV  
DISLTPVPAEDNENSEYFGQNNVNTDTATNATKATATWTKGTPVPERKALY  
CNLWAGFTDRLSGLVEIGATGMEKQFVWSMNFHTKGTDEKRRKFGRTSGY  
VIGSATPTPKDDLETFAPFCHLPARDKCFIAHNSRRTYGGTLFFKSHHTLPQNTYRL  
GRKFSSEATKEPRTPLADVOVSHSNDENERTYTLPSSEGSMECAAGCIG  
LDLPVLSNHPKFTPTIPOMKEMTYVSGNSFESSDGRGSLGHLNLSIPVAK  
FVGDIDISTYTDLSGFVSDVRRNNSQATLWDSKIRGNLSRQAFILRGSN  
NYYNSNCELFGYAMLRGSSRNYNVDGKLRP"  
5026..7812  
/gene="CP0303"  
5026..7812  
/gene="CP0303"  
/note="similar to GP:4376729; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="polymorphic membrane protein G family"  
/protein\_id="AA38160.1"  
/db\_xref="GI:7189230"  
/translation="MKSFSEVLVSLTACFTSCSYFPAATAENIGSPDSFDSNTG  
TYTPKNTTGIDITLIGDITLQNLGDSALTKCFSDTTSLSFAGYVLSFLNKS  
SAGGALSVTTDKNLSTGFSLLPLFASVTLTPSGKAVCGGGLPDDNNTLIF  
KODYCENGGATSTKNLSKNSGTSIFENKSAAGKAGCAICAGTDTWNTAPRT  
LFSNNITAEAGALINSNGCTTIGTNSVISENSVIRATNAGALSADADYITSGNS  
VITSGNAVANGGALIAKRLTLASGGGGISFSSNNIVOGTTAGNGGALISLAGECSL  
SAEAGDTFENGNAIVATPQTRNSIDIGSTAKITNLRAISGHSIFFYDPAENSLA  
DSTDITLNRADAGNSDYSGSIVFSGEKLSEDEAVADVLLSLKQVYLLAGNVYL  
KRGTLDTRKGTOTAGSSVIMADGTLTKASEEVLTLGSLIPVDSLGEKRYVIAASA  
ASKNVALSGPILLDNONGNAYENHDKTDFEVLPSAAGTTPDPAVYVATPRT  
HYVQGTWMTWTDPASTPKRTATLANTGATLPSVPSAGGLVNSLMSGSSDQIA  
IOGVIRSATLISDSGFPNAGYANLDRDKKREKRYRSGGIVAGAQVCSNL  
ISFRCQLFESDQDLVAKNHTDTVGAFTQHTTCSGTLGCLLPLKPSWHSKPLV  
LEQQLAVSHVNDLKTRTAYPEVKSWMGNANMAGASHSGLPYLKFQFDYVAYI  
KLNLTYIRQDSFSEKTEGRSFDNSLFINSLDLYGVFEKSCDNFSYDLTLSTYVD  
LIRNDPKCTALVYSGASWETVANNLRAQLQVRAQSHVAFSPMFVLQGFVEVNGS  
SRIVNDLGGKRFQ"  
complement(8170..8292)

```

CDS
/gene="CP0304"
/complement(8170..8292)
/note="CP0304"
/note="hypothetical protein; identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA38161.1"
/db_xref="GI:7189231"
/translation="MDSSRWLNKLRKRLDQNDIKYCKSYCLRGISQLLGMW"
8329..9444
/gene="CP0305"
8329..9444
/gene="CP0305"
/note="conserved hypothetical protein; identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA38162.1"
/db_xref="GI:7189232"
/translation="MEMSPPEQPEQCHFDVVSFLPESLITRARSDFEGRITVEQM RVEDDAIRNLIKRQTEAGLFTFDGEFRYSMDPDMGFHGVDRSDNDEIGY LKDIISVSKHPIEFHFEVKEFKENAKAKOTIPSPQGFHEMFAENLNTKRFPT NOELIDIVFYROYIODLYAAGRNOLDQCAMCRILDRASWGVDSHDLQEL EOLFMINLVKMRPEDEFLVSLHYCRGDQAEFSPRAYSIEPLFAKIDVSIHY WALDXYSGGAELPAAVYSGERHVLGLISSHSCIEDRDVSKTYEASVLEPLERLS LSPCGFASCEGDRHMTEDDMKIAFVKELAEIKW"
complement(9785..12571)
/gene="CP0306"
complement(9785..12571)
/note="similar to GP:4376731; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="polymorphic membrane protein G family"
/protein_id="AA38163.1"
/db_xref="GI:7189233"
/translation="MKSSLWFLISSSLAPLSINSAFAVIEINLGPNSGEGCT YTPAOTTNADGTYLNTGDVSTNNAGSPALASCKERTGNLSFGHQYFLLONI DAGANCTFTTANAKLSFSFSTYLSLTQTNNAKTGALSGAGSISQNSCYEQ NPSNDNGALQSSISLSLNPNTLFAKNKATGALSGGALSTGNTNNNSAFSENT AANNQAIYEAASFISNKAISFINNSVATATGAGAIYCSSTAPKPLTSDNGE LNFIGNTAITSGAIYTDNLVSSGGPTLKNNSAIDTAPAGGAIADSGSLISA LGSDITPEGNTVYKGAASSQTTNNTNIGNTAKIVQARSAGNTITTYDPTTISIT AALSDALNNGDPLAGNPRATOGTIVISGKELSAEAEADNLKSTIQOPLTAGGOLS LKSGYTLVAKSFQSFGSTLMDAGTLETAADITNNVNLNLSLKEKTKATLKATQ ASQTVLISGLSLVDPGNYEDVSWNNPDVFCQLTADDPANIHITDLADPLEKN PIHWGQGNMALMOEDTATKSKATLTMTKGTYNPERGRLVANTLMSFVDRS IQOLVATKYROSQETRGIMCEGTSNPFHDSGTINGFPHISAGYVVGATTLASNL ITAPFOLFCKDRDHFINKRASAAYASLHIOALATLSSPLIRIPGESQDPVLEF AOISYIYHDSFKERNITLVKSPSDLVINVAITGTFERFSENEASVYAVIYAD VYKRNPDCTTALINNTSWKTGTNLNLRQAGVIRAGIFYAFSPNLEVTNLSMEITGS SRSYNDLGGKRF"
BASE COUNT 3580 a 2439 c 2890 g 3767 t
ORIGIN
Query Match 17.7%; Score 531.8; DB 1; Length 12676;
Best Local Similarity 53.0%; Pired. No. 5e-104;
Matches 1416; Conservative 0; Mismatches 1182; Indels 75; Gaps 10;

```

```

QY 367 cgggaactctctatgttccaaacgctgagtcgaggaactgtagcagggctgctgttaa 426
Db 5298 GGGGTACTCACTTCTTTTAAATATTAAGTCTAGTGTGAGGGCGACACTTCTGT 5357
QY 427 cagcagctggtgataataatccacacgcttataaggtttcttcgtatctttatgc 486
Db 5358 TACAAGTGAATAAATATGTGCTTACAGGATTTTGGAGTCTTCTTACGGCCCC 5417
QY 487 gtccctcggaagttcgaatactaccgcgaagagccgttagtgcgtctacaggtgctt 546
Db 5418 ATCATCGGTATTCACACCCCCCTCAGGAAAAGTGCAGTTAAATGAGAGGGATC---T 5474
QY 547 gaattgacaaaatgctcagttgctcttcgaacaaactttcaacgataatgcg 606
Db 5475 TACATTGTGATAACAAATGAGACTATTTTATTAACAAGATTACTGTGAGAAAATGGCG 5534
QY 607 tgcataccgcgaanaactttcattacaggaactacaatgctcagctctgtttctga 666
Db 5535 AGCCATTTCACCAAGAACTTTCTTGAAAAACACACGCGGATCGATTCTTTGAAAG 5594
QY 667 aaata-----ccctccaaagaagcgagccattcagacttcgagcccttac 717
Db 5595 GAATTAATCAGGCAACAGGAAAAAAGTGGGCTATTGTGCTACTGTACTGTAG 5654
QY 718 cattactggaacccaagggggaagtcctttctgcaaatcttctcgatctgagc 777
Db 5655 TATTACAAATATATACGGCTCTTACCCCTCTTCGACAAATATTGCTGAGAGTG 5714
QY 778 tgcatttttaccagagccctcggtgactttcttaataatgttaagttcccttattga 837
Db 5715 AGCTATTAATATGACACAGAACTGTACAAATTACAGGAAATAGCTCTTGTATTTCTGA 5774
QY 838 caataagttacagagcgagctcctcaacaacggygataatgtaagaggtgactatcg 897
Db 5775 AAATAGTGTGAC---AGCAGCCGACAGAAATGAGAGAGCTCTTCTGGAGATCGCG--- 5827
QY 898 tggttataaactcagtaactaagttaccctcactcgtgaatcagatgttaactcct 957
Db 5828 -----ATGTTCAATATCTTGGAATCAGAGTCACTT 5861
QY 958 cagcaacaatacgcacacacagcggaagagctatctatggaanaacgcgaactcgc 1017
Db 5862 CTCGAGAAACCAAGCTGTAGCTAATAGCGCGACCACTTATGCTAAGAACTTACACTGGC 5921
QY 1018 ttccggaagact-----tacccattcagtagaataatggtcgaatggaggtacagctcc 1071
Db 5922 TTCGGGGGGGGGGGGGTATCTCTTTCTTCAAAATATATAGCCAAAGGTACACATCG 5981
QY 1072 taaggttgagccataagctatcgaagataagtgaggaaattgattatccgcgataatg 1131
Db 5982 TAATGTGAGCCATTTCTTACTGTGACCTGGAGAGTGTGATCTTTTACGAGAAAGCAGG 6041
QY 1132 tgacattgctttttaggaatacagtaactcctactactcc-----tggaagaaataag 1188
Db 6042 GGACATTACTTTCATGGAATGCGCATTTGTTGCAACTACACCAAACTACGAAAGAA 6101
QY 1189 tagtatcgacttaagaaagagtgcaaaatgacagcttgcgttcgtctgtgtagagc 1248
Db 6102 TTCTATTGACATGATCTACTGCAAAAGATCAGAAATTTAGTGAATATCTGGGCATAG 6161
QY 1249 catctactctatgtatccataactacaggaatccatcccaacagttaaagtgtccttaa 1308
Db 6162 CATCTTTTTCACGATCCGATTTACTGCTAATACGGCTCGGATTTCTACAGTAATTTAA 6221
QY 1309 agttaatgagactccgagagattctgcactacaatatagaggaacatcatcctcaag 1368
Db 6222 TCTCAATTAAGGCTGATGAGGAATATAGTACAGATTATAGTGGTCGATGTCTTTTCTGG 6281
QY 1369 agaaagttatcagagaaagagcgacagattctaaatcttacttcogaaagtactaca 1428
Db 6282 TGAAGAGCTCTGTGAAGATGAAGCAAAAGTTGACAGACAACTCACTTACGCTGAAGCA 6341
QY 1429 gcctgttaactcttcagaggttactatctttaaacatgagtgagctcgcagactca 1488

```

Db	6342	GGCGTGAACCTAAGCTGAGAAATTAGTACTTAAACGGGGGTACTCTTGAAACGA	6401
QY	1489	ggcattcaactcaacagcgagattcgtctcgaaatgagagtaggaactactcgaac	1548
Db	6402	AGGGCTTACTAGACCGCGGGTTCCTGTTATTATGATGCGCGGCACAACTTAAAGC	6461
QY	1549	---tgcataactagcaacataaacaattgycataacatcagtcactgactcgtagcggc	1605
Db	6462	AAGTACAGAGGAGGTCACTTAAACAGGCTTTCCATTCTCTGAGACTCTTTAGCGAGG	6521
QY	1606	aagaaggcaaaaatagaaaacaaagctacgylcaaaaactcyacttactcgaacat	1665
Db	6522	TAAGAAATTTTAAATTTCTGCTTCTTGCGACAGATAAAAAATGAGCCCTTAGTGTCAT	6581
QY	1666	caattatggaaccgcagcgacggtttttagaaatcagtttagaaatccctcagc	1725
Db	6582	TCTTCTTTGGATTAACCAAGGAAATGCTTATGAAATTCAGACTTAGGAAAACTCAAG	6641
QY	1726	ctacgacacttaagagctcaaaagcttcctggaactgtaaaagacccgcgttaaccga	1785
Db	6642	CTTTCAATTGTGACGCTCTGCTGCGGTACTGCAACAAGTAAAGATTCACGCGT	6701
QY	1786	tccataatggtgtagaatccattccattcagcgtctcagggaaacttggggccaattgtt	1845
Db	6702	TCTTACACTACCAACTCTTACGCACTATGGGTCTCAGGACTTGGGGATGACTTGGGT	6761
QY	1846	gg-----ggaacagggtcttcaagatgtaaccttaactgaactgaactaaacg	1893
Db	6762	TGATGATATCCCGACAGACTCCCAAAAGACTAAGACAGCATTAAGCTTGGACCAATACG	6821
QY	1894	ctatactcaatcccgagcgatcggtctcttagtccttaatagctttagatgcat	1953
Db	6822	CTACCTTCCGATTCCTAGCGCTCAGGACCTTAGTACTTAATAGCCTTGGGGATCTTT	6881
QY	1954	tatagataagcctctccattactcctaagagactcgaaacgaaaggtttgcaggga	2013
Db	6882	TTTCAGACTCCAAACGATTCAAGTGTCAATAGAGAAAGTCTTGACTCTTTGTTCCA	6941
QY	2014	ccgtgcttcttggtgctgcggtatctaaactctctccataagagatgatacaaaacg	2073
Db	6942	TGCGAGCTTCTGGGCTCGCGGAGTCGCCAATTTCTTAGTAAAGATAAAGAGGGGAAA	7001
QY	2074	acggcggtctgcgaattggtggcggtatgctcagtagaggaacctccaaactgtct	2133
Db	7002	ACGCAATACCGTATATAAATCTGCTGGATATGCTATCGGAGGCGCAAGCCAACTGTCT	7061
QY	2134	agataagattcttagtgcgtcatctttgtccagcctcttggaaagatagagactactgt	2193
Db	7062	TGAAAACTTATTTAGCTTTCGCTTTGCCAAGCTTTGGAGGAAATTAAGATTCTTAGT	7121
QY	2194	agctaaagatcaaggtcagctctcagaggaagaaactctctatccagcaaacgaacct	2253
Db	7122	CGCTAAATATATATCTATACCTATGAGGAGGCTTCTATATCAACAAATTAACGAATG	7181
QY	2254	tatctctcttcctgcgaactcagcgctgttcgtgtctatgttctccagagatcc	2313
Db	7182	TATGCGGTTCATAGGTGTCTCTTAGATAAA---CTTCTGCTCTTGGAGTCAATAAC	7238
QY	2314	tgtctctttcaggaacaccttagctacacccaacagctaaagaaacgtctgaaaccaa	2373
Db	7239	CCTGCTTTTAAAGGCGACGCTCGCTTATAGCCACGTCAATATGATCTGTAACAAATTA	7298
QY	2374	tacaacatctccactctttaaaggagcgctggggaatgtagttcgtcttagaatctg	2433
Db	7299	TATCGCGATACCTTAGGTGAAGGTTCTTGGGGAAATAAGCTTTTAACATGATGTTGG	7358
QY	2434	tgsaagagctccgaattgtcttagaagaaggtgtctcattatgttagcagtcacatgcc	2493
Db	7359	AGCTTCTTCATCTTATCTTATCTGAATAACCGATCGATTTGTTTGAATACCTAAGCTCA	7418
QY	2494	gaatttcagttgtctcagcatcaggaaggttttaaagaacgggaacgagctgc	2553

Db	7419	CANCTGATATCTGACACATATACCTGACGACACCTTCTCGGAAAGATACAGAGAA	7478
Oy	2554	tgaaattggaaatgcgcgcgtctgtgaaacctgtgcttaccatcgtggatccgaattgataa	2613
Db	7479	ATCTTTTATGACAGCAACCTCTCTCAATTTATCTTTGCCATATAGGGGTGAAGTTTGANA	7538
Oy	2614	ggaaatgcagacgccaagaatgcaacggtacatctcaactctctgtgtatacgtgattgt	2673
Db	7539	GTCTCTGATATGTAATAGACTTTTCTTATGTAGTCGACTTATCTTAATGCTTGATCTTAT	7598
Oy	2674	tcgtatgaaccgcgcgttgcgaacaacgcgcgaattgacgctggtatcttggaaaacct	2733
Db	7599	CCGCATGATATCCCAATATGCACTACAGACACTGTATATCAGCGGAGCCTCTTGGGAANCTTA	7658
Oy	2734	cgtgtaacgaattggccaagaacgaacttaagctctcgtgcagggaaaccattttgcttaa	2793
Db	7659	TGCCATTAACCTTACACGACAGCGCTTTCGAAGTCGCGTGCAGCGACACTACGCCCTCTC	7718
Oy	2794	ctcaaatattgaagccttgcgaacttcttcttggatgacgtgggtcatctgcgaatta	2853
Db	7719	TCCATATGTTTAATGCTCGCCACGTTTGTCTTGAAGTTCTGTGATCCTCACGGATTTA	7778
Oy	2854	caatgtgaacttaaggacaaataccaattcga	2886
Db	7779	TAAATGATCTTGGGGGTAAAGTTCCAAATCTTA	7811
RESULT	9		
AP002546/c			
LOCUS	300600 bp	DNA	BCT
DEFINITION	Chlamydomophila pneumoniae genomic DNA, complete sequence, section 2/4.		07-JUL-2000
ACCESSION	AP002546	AB033782	AB033783
	AB033782	AB033783	AB033784
	AB033802	AB033803	AB033804
	AB033804	AB033805	AB033807
	AB033808	AB033810	AB033811
	AB033812	AB033813	AB033814
	AB033815	AB033816	AB033817
	AB033818	AB033819	AB033820
	AB033821	AB033822	AB033823
	AB033824	AB033825	AB033826
	AB033827	AB033828	AB033829
	AB033830	AB033831	AB033832
	AB033833	AB033834	AB033835
	AB033836	AB033837	AB033838
	AB033839	AB033840	AB033841
	AB033842	AB033843	AB033844
	AB033845	AB033846	AB033847
	AB033848	AB033849	AB033850
	AB033851	AB033852	AB033853
	AB033854	AB033855	AB033856
	AB033857	AB033858	AB033859
	AB033860	AB033861	AB033862
	AB033863	AB033864	AB033865
	AB033866	AB033867	AB033868
	AB033869	AB033870	AB033871
	AB033872	AB033873	AB033874
	AB033875	AB033876	AB033877
	AB033878	AB033879	AB033880
	AB033881	AB033882	AB033883
	AB033884	AB033885	AB033886
	AB033887	AB033888	AB033889
	AB033890	AB033891	AB033892
	AB033893	AB033894	AB033895
	AB033896	AB033897	AB033898
	AB033899	AB033900	AB033901
	AB033902	AB033903	AB033904
	AB033905	AB033906	AB033907
	AB033908	AB033909	AB033910
	AB033911	AB033912	AB033913
	AB033914	AB033915	AB033916
	AB033917	AB033918	AB033919
	AB033920	AB033921	AB033922
	AB033923	AB033924	AB033925
	AB033926	AB033927	AB033928
	AB033929	AB033930	AB033931
	AB033932	AB033933	AB033934
	AB033935	AB033936	AB033937
	AB033938	AB033939	AB033940
	AB033941	AB033942	AB033943
	AB033944	AB033945	AB033946
	AB033947	AB033948	AB033949
	AB033950	AB033951	AB033952
	AB033953	AB033954	AB033955
	AB033956	AB033957	AB033958
	AB033959	AB033960	AB033961
	AB033962	AB033963	AB033964
	AB033965	AB033966	AB033967
	AB033968	AB033969	AB033970
	AB033971	AB033972	AB033973
	AB033974	AB033975	AB033976
	AB033977	AB033978	AB033979
	AB033980	AB033981	AB033982
	AB033983	AB033984	AB033985
	AB033986	AB033987	AB033988
	AB033989	AB033990	AB033991
	AB033992	AB033993	AB033994
	AB033995	AB033996	AB033997
	AB033998	AB	



Query Match	17.7%;	Score 531.8;	DB 1;	Length 300600;
Best Local Similarity	53.0%;	Pred. No. 3.8e-104;		
Matches 1416; Conservative	0;	Mismatches 1182;	Indels 75;	Gaps 10;

Dh	206804	TAATGGTGGAGCCATTCTTCTATACGCGAGCGTGGAGAGTGTAGTCTTTCCACGACAGACGG	206745
Oy	1132	tgacatgctctcttttaggaaataagcactctctactacc---tggacgaataag	1188
Dh	206744	GGACATTAACCTTCATTAAGGGAATGGCAATGTGGCAACTACCCACCAACTCGAAGAAAGAAA	206685
Oy	1189	tagatcgacttaagaaacagtgcaaaagatgcaagccttgcgtctgctgctgtagagc	1248
Dh	206684	TTCTATTGACATAGAGACTACTGCGAAAGATCCAGCAATTTACGTGCAATATCTGGGCATAC	206625
Oy	1249	catctactctcatatgccataactacacagatcatccacaacagttacagaigtcthaa	1308
Dh	206624	CATCTTTTCTACACATCCGATTACTGCTAATACCGCTGCGGATTTCTACAACTCTTTAA	206565
Oy	1309	agttaatagacatccggcagatctctgacatacaataacggggaaacatactctacag	1368
Dh	206564	TCTCAATTAAGGCTATGTCAGTAATTAATACAGATTATAGTGGGTCGATTTTCTGG	206505
Oy	1359	agaaagttatcagagacagagcgccagattcttaaaaacttactctcgagacttaca	1428
Dh	206504	TGAAAAGCTCTCTAGATGAGCAAAAGTTGCGACAACTCAGCTCTACGGTGAAAGCA	206445
Oy	1429	ggcgttaactcttcagagagttactctatctttaaacatgagagtgaactctcgagactca	1488
Dh	206444	GCCTGTAACTCTTAACCTCAGGAAATTTAGTACTTAAACGGTGGTCCCTCTCATACGAA	206385
Oy	1489	ggcattcactcaaacaggcagatctcgctccgaatgagcttagaactacttagaac	1548
Dh	206384	AGGCTTAACTCAGACCGGGTTCTCTGTATTATATGATGCGGGCACAACGTTAAAGC	206325
Oy	1549	---tgctgatactcgacccataaacaattggtcattacaatcagtctctatagacggtgc	1605
Dh	206324	AAGTACAGAGAGAGTCTTTAAACAGTCTTTCATCTCTGTAGACTCTTTAGCGCAGGG	206265
Oy	1606	aagaagaagcaaaaatagaaaccaagtaagtcataaaaatctgacttacttagaacat	1665
Dh	206264	TAAAGAAAGTTGTATTGCTGCTCTTGCGACGAAGTAAATGTAGCCCTTAGTGTGCCAT	206205
Oy	1666	caattatltggaccgcgaagggcacgttttatgaaaatcatagitttaagaaatccctgac	1725
Dh	206204	TCTTCTTTTGGATTAACCAAGGGAATGCTTTGAAATACAGACTTTAGAAATCTCAAGA	206145
Oy	1726	ctaaagacatcttagagctcaaaagctcttgaaactgtaaacagacacggcagtgactcaga	1785
Dh	206144	CTTTTCAATTGTGCAGCTCTCTGCTGGGTACTGCAACACTACAGATTTCCACGGT	206085
Oy	1786	tccataatgggtagaagaattccatacgtactacaggtactacaggaactgggcccacatttg	1845
Dh	206084	TCTTACAGTACCACTCCTACGACATATGGGTATCAAGGATACCTTGGGGATATGACTTGGT	206025
Oy	1846	gg-----ggacagagggtctctacagatgcaacatccaactggaactaaacttgg	1893
Dh	206024	TGATGATACCGCAAGCACTCCAAAGATTAAGACAGCACCTTGTGGACCAATACAG	205965
Oy	1894	ctatactcctaactccgaaggtatccggtctttagtcccaataagttatgagatgatt	1953
Dh	205964	CTACCTTCCGAATCTTAAGCGCTCAAGGACCTTTTAGTTCTTAATAGCCTTTGGGGATCTTT	205905
Oy	1954	tatagataatagctctctcatatctcttaatggaactgcaaacagaaaggttgcagaggaga	2013
Dh	205904	TTTCAGACATCCAAACGATTCGAAGGTGCATAGAGAGAAAGCTTTCACCTTTGTTACAG	205845
Oy	2014	ccggtgcttttgggtgctggtatctactctctccataaagatagtagtcaaaaacag	2073
Dh	205844	TCGAGGCTCTGGGCTCGGGAGTCCGCCAATTTCTTAGATAAAGATPAAGAAAGGAAAA	205785
Oy	2074	acggagggttccgcatcttgagtgcggttatgcatagagggaaactcatatcattgctc	2133
Dh	205784	ACGCAAAATACCGTATAAATCTGCTGATATGCTATGAGAGGTGCAGCGCAAACTGTTC	205725
Oy	2134	agataagatctcttagtgcgtcattttgtcgaagctctcttgaagaagatagagactattgt	2193
Dh	205724	TGAATAACTTATTAATGCTTCTGCTCTTTGCCAATCTTTGGTAGGCAATPAAGATTTTCTAGT	205665

RESULT	10								
LOCUS	AE001628/c								
DEFINITION	Chlamydia pneumoniae	section 44	of 103	of the complete genome.					
ACCESSION	AE001628	AE001363							
VERSION	AE001628.1	GI:4376730							
KEYWORDS									
SOURCE									
ORGANISM	Chlamydia pneumoniae.								
REFERENCE	Chlamydia pneumoniae.								
AUTHORS	Bacteria: Chlamydiales; Chlamydiaceae: Chlamydia.								
	1 (bases 1 to 10757)								
	Kalman, S., Mitchell, W.,	Marathe, R., Lammel, C., Fan, J., Olinger, L.,							
	Grimwood, J., Davis, R.W. and Stephens, R.S.								
	Comparative Genomes of Chlamydia pneumoniae and C. trachomatis								
	Unpublished								
	2 (bases 1 to 10757)								
	Kalman, S., Mitchell, W.,	Marathe, R., Lammel, C., Fan, J., Olinger, L.,							
	Grimwood, J., Davis, R.W. and Stephens, R.S.								
	Direct Submission								
	Submitted (01-DEC-1998)	Program in Infectious Diseases, University							
	of California, 235 Earl Warren Hall, Berkeley, CA 94720, USA								
	Location/Qualifiers								
	1. 10757								

/organism="Chlamydomophila pneumoniae"  
 /strain="CML029"  
 /db\_xref="taxon:83558"  
 106..2892  
 /gene="pmp\_9"  
 /note="Cp00447"  
 106..2892  
 /gene="pmp\_9"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Polymorphic Outer Membrane Protein G/I Family"  
 /protein\_id="AA018591.1"  
 /db\_xref="GI:4376731"  
 /translation="MMSLSLHWFISSSLALPLSNFAFAVEINLIGTNSFSGG  
 YTPADPTNADGDIYKLTLDSTNGSPALTALASCFKRTKSLSGQGYPLION  
 DAAPFNTNANKLTLSTSGSYSLIOTLTATTTGAIKSTGACISQNSYCYFG  
 NFNNGDGLAOGSITLSLSPNLTTPKKNATOGGALYSTGSIITNLTNSAFSESN  
 AANNGDIATGASPTISSKKAISFINNSVTATSAVGAIYCSSTAPKVLVLSNG  
 LNFICGATISGAIYTDVLVLSGGPTLFKNNMSAIDTAPLAGAIALNDGSLSL  
 LGGDITFESNTVYKAGSSSTPLTNRNSINGTNNAKIVOLRASQNTFYEDPITSI  
 LSGVTLVKSLSVPSGQVYEDVGMNDPOVSCITLRDDPANIHTDLADPEK  
 ASQVTLVKSLSVPSGQVYEDVGMNDPOVSCITLRDDPANIHTDLADPEK  
 PLHWGQGVWALSMQEDTKERKATLYTKTGVNPERRGITVAANTLMSGVYR  
 IOIVATYRQSGQERTRIGCBISINFEHDSYKIKGGRHSAGVYATGTLISDN  
 ITAAICQLFGKRGRIINNRSASAAASHLQHLATLSPSLRYLPSESEPOVPLD  
 AOSTISHTDSEFKERTYVLPAPKSESSWDGCLAEIASSLPHTALHSLFHAPEPT  
 VEASTIHDSNFERNTLVRSFSDGLINVSPIGITFERSRNRASAEVATYVAT  
 YRRNPDCTTLALINNTSWKTGTNTLSHQAGIGRGIYFASPNLEVTNSMETRG  
 SSVYADADGGRKQF"  
 complement(3233..4348)  
 /gene="yxj6\_2"  
 /note="Cp00448"  
 complement(3233..4348)  
 /gene="yxj6\_2"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Hypothetical Protein"  
 /protein\_id="AA018592.1"  
 /db\_xref="GI:4376732"  
 /translation="MEMSPQPEQEQCFDVGVSFLRPESITLRASDFEGRIYVEO  
 RYVDAALIRNLIKQTKAGLTFEDGFERRYSWPDFGMWGHGVDRSDNDPEIGY  
 LKQELISVSKHPIIFEFYVTEFGNAKAKOTIPSPQFPHMEIAPNTNKRFTYT  
 NQELIDIVSEYFVYIODIVAAGRNQIOLDCAACRLDIRBSMYVSDHDKLOEL  
 EQFLMINLVKDKDEPLFSLHCRGDYQAEFRSRAIYSIEPLFAKIDVSYHY  
 WALDQSGSGGEPLIYASGKRIKGLISSNHSIEDRDVVSRIYEAASYIPLERLS  
 LSPQGRASCEBDRHREEDCKMKIAFVKAEIKEW"  
 complement(4865..6769)  
 /gene="pmp\_10.1"  
 /note="Cp00449"  
 complement(4865..7650)  
 /note="Polymorphic Outer Membrane Protein G Family;  
 contains frame shift"  
 complement(6715..7650)  
 /gene="pmp\_10.2"  
 /note="Cp00450"  
 7813..10599  
 /gene="pmp\_11"  
 /note="Cp00451"  
 7813..10599  
 /gene="pmp\_11"  
 /codon\_start=1  
 /transl\_table=11  
 /product="Polymorphic Outer Membrane Protein G Family"  
 /protein\_id="AA018593.1"  
 /db\_xref="GI:4376733"  
 /translation="KKTSPWVYSSVLAFCISLONLAEELLSPDSENGINDSGTF  
 TKTSTATYSLICGVDFYIPBGVGLPISDCFCQOTDNLTFILONGHSLFGFIDATHA  
 GAASTTAKRNLTFEGSLSLSDSPSTIVTIGQTLISAGAVINNTYRKLVAANPS  
 IADGGAIKASPLLTGSDALSNSSSTKGAJATTAGARIANNTGYRPLTASNS  
 TSGGAIIDEGTSLINNNKFLVEGNAAKTGGALCNTASGSGPELIINNKTLITASN  
 VAEETGGAHAKRLISGSGFEPLRNVSATTPGCAISIDASELSLSATGQITF  
 VNTLTLTGSTDTPRRNNINISNCKFTDLRAAKNHTLFFDPTIESTSDVLIINN

GSAGALNPYQCTIFSGEFTLPADELKVADNLKSSFTQPVSLSGKLLLOKQVTESTS  
 FSOEAGSLGMDSGTSTETAGSITITMGNVSLGKQVSLSTAKASNKVIVSGK  
 LNLIDIEGNIENSHMEFSDQLSLKIVDADVDNVDISLIPVADPVSSEFQ  
 OMVNTMTDITNTEKATWTKTGEVSPERKSLVNTLITAGVFTDRISLOLVEIG  
 ATGHEHKGFPVSSMTNFKTKGDNREKRGFTSGGYIGSANTPRNDLTPFACHL  
 FARDKDCIANNNSKTYGCTFFPKSHSLQONLIRLDRASFSALEPFREIPLAL  
 DVQVFSHSDNRMEHTHTSLPESGSWSNEQIAGIGIDLEFVLSNPHLEKFIPOK  
 KVEWYVQNSFEFSSSDGRGFSIGRLNLSIPVAKVOODIDISYDIDSGEFVSD  
 VYRNPOSTATLIVMSPDSMKIRGSLRQAEFLRLGNSNNYVNSNCELEGHYAMELRGS  
 SRNINVDTGTRLEF

BASE COUNT 3228 a 2459 c 2065 g 3005 t  
 ORIGIN

Query Match 17.7%; Score 531.2; DB 1; Length 10757;  
 Best Local Similarity 53.0%; Pred. No. 6.8e-104;  
 Matches 1415; Conservative 0; Mismatches 1183; Indels 74; Gaps 10;

QY 247 tgcctgaactactctactctttaaggaaatgtcaccttagaataatctctggaacagg 306  
 DB 7435 TACTGGAAATAGACTACTCTGACAGAGATATTAACCTGCAAAACCTTGGGGATTCCG- 7435  
 QY 307 caccgaatcacaaaagctgtttaacacacactaaggcgattgacttcacaggtaa 366  
 DB 7436 --CAGCTTTACGAAGGGTGTCTTTCTGACACTACGGAATCTTTAGCTTTGCCGTTA 7379  
 QY 367 cgggaactctctatgttccaaaacggtggaatgcaaggactgacaggggctgtctgttaa 426  
 DB 7378 GGGGACTACACTCTTTCTTTTAAATATTAAGTCTAGTCTGAAGGGCGACACTTCTGT 7319  
 QY 427 cagcgcggtgtagataatctacacacgtttaagggttcttgcctctcttaattgc 486  
 DB 7318 TACAACCTGATAAAAATCTGCTCAACAGAGATTTGAGCTCTTACTTTCTTGAACGCGCCC 7259  
 QY 487 gtccctcgaagctcgataacacccgcaaggacggttagctgtcctacaggttagctt 546  
 DB 7258 ATCATCGGTAATCACAAACCCCTCAGGAAAGTGCAGTTAATGTGGAGGGATC---T 7202  
 QY 547 gaggttgcaaaaaatgtcagtttgccttcagcaaaaaacttcaacggaataatgacg 606  
 DB 7201 TACATTGTGATTAACATGGAATCTTTTATTTAAACAGATTAAGTGTGAGAAATATGGCG 7142  
 QY 607 tgcctcacccgcaaaaacttcttaacagaggaactaaatgtaagctctgtttcttga 666  
 DB 7141 ACCCATTTCTACCAAGAAATCTTTGAAAAACAGCACGGATGATCTTTTGAAG 7082  
 QY 667 aataa-----cctcccaagaagaagcgagccattcagacttccgaatgacctaac 717  
 DB 7081 GAATAATCGAGCGCAACAGGAAAGAGTGGGCTATTTTGTCTAGTGTACTGTAGA 7022  
 QY 718 cacttctggaacacaaagggaagctcttcttctgaacaaacttcttcggattctgagac 777  
 DB 7021 TATTTCAAATATACGGCTCTTACCCCTCTTCGAACAATTTGCTGAAGCTGCAAGGTGG 6962  
 QY 778 tgcatttttcaagaagcctcggtgactatttctaataatgtaagtttcccttaattga 837  
 DB 6961 AGCTTAATATACACAGGAAGTGTACATTTACAGGAAATACGCTCTTTGTATTTTCTGA 6902  
 QY 838 caataaggtcacagagagcgtctctcaacaaagggtgataatgtaagaggtgtctatctg 897  
 DB 6901 AAATAGTGTGAC---AGCAGCCGCAAGAAATGAGAGAGCTTTCTCGAGATGCGG--- 6847  
 QY 898 tgcctataaactagtaacatacaaggtcacctcactggaataatcagaatttactctt 957  
 DB 6848 -----ATGTTACATATCTGGGAATCAAGTGTACTTT 6815  
 QY 938 cagcaacaatatcatcgacacagcggaagactatctatgtgaaaaagctcgaaatgac 1017  
 DB 6814 CTCAGGAACACCAAGCTGTATGTATGTGGCGAGCCATTTTGTCTAGAGACTTACACTGGC 6755  
 QY 1018 ttcggaggaact-----taacctatcagtagaataatgtaatgtaaggtatcagctcct 1072  
 DB 6754 TTCCGGGGGGGGGGGTATCTCTTTCTAACAATATATGTCCAAAGTGTCCACTGACAGT 6695

QY 1073 aagggtgagccatagatctatcgaaagatagtggggaattgatttcccccagataggt 1132  
 DB 6694 AATGTTGAGACCATTTCTATTAAGTGCACCTGGAGAGTGTAGCTTTTCCAGCAAGACAGG 6635  
 QY 1133 gacattgtcttctttaggaataacagctcacttactactcc---tggagcaatagaagt 1189  
 DB 6634 GACATTACTCTCAAGGAATGCCATTTGTTCACTACACACCAACAAACTCAAAAAAGAAAT 6575  
 QY 1190 agtatcgacttaggaagagagtgcaagaatgacagcttggttcgttcgcgcgtgtgagacc 1249  
 DB 6574 TCTATTGACATAGGATCTACTACGAAATGACGAATTTACGTGCAATATCTGGCATAGC 6515  
 QY 1250 atctactctatgatccataactacaagatcatccacacagttacagatgtctttaa 1309  
 DB 6514 ATCTTTTCTAGACATCGCATTAAGTCTTAATAGCGGTGGGATTTCTACAGATTAATTAAT 6455  
 QY 1310 gttatgagactcggcaagatctgcactacaatatcacaggaacatcattccagga 1369  
 DB 6454 CTCAATTAAGGCTGATGACAGTAAATGTAAGATTAATAGTGGGTGATTTTCTGT 6395  
 QY 1370 gaaaggttatcagagacagagcgacagatctctaaataacttactcgagactacag 1429  
 DB 6394 GAAAAGCTCTGAGAGATGAAGCAAAATTTGACAGCAACCTCATTCTTACGCTGACAGAG 6335  
 QY 1430 cctgttaactcttcagaggaactctatctttaaacaaggagtgactgtcagactcag 1489  
 DB 6334 CCTGTAACTTAAGTGAAGAAATTAAGTAACTTAAGTAACTGAGTCACTCGATACGAA 6275  
 QY 1490 gcatcactcaacagagcagatctcgtctcgaaatgagcgtgaagaaactacttagaac- 1548  
 DB 6274 GCTTTACTCAACCGCGGGTCTCTGTATTATGAGTGGGGGACACACGTTAAACCA 6215  
 QY 1549 --tgtctactaagcaccataaacaattggtcttaataacatcagatctctataggttga 1606  
 DB 6214 AGTAAAGAGAGGTACTTTAAACAGTCTTCCATTTCTGTAGACTCTTTAGGCGAGGCT 6155  
 QY 1607 aagaagcgaataatagaacccaagaactcagcaaaaaacttgaacttacttgaacacatc 1666  
 DB 6154 AAGAAAGTTGTAATGTGCTTCTGACAGAAAGTAAATATGAGCCCTTAAGTGTCCGATT 6095  
 QY 1667 acttattgagcccgacagcggttcttaagaataacttaagaaactcactagctc 1726  
 DB 6094 CTTCTTTGGAATACCAAGGAGATGCTTATGAAATTCAGACTTATGAAAACTCAACAGC 6035  
 QY 1727 taagacatcttagagctcaaaagctctgaaactgaaacagccagctgactcagat 1786  
 DB 6034 TTTTATTTGTGACACTCTGTGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 5975  
 QY 1787 cctataatggtgagaaatccatgaagctacaggaacttggggcccaattgttgg 1846  
 DB 5974 CTTACAGTAGCAACTCTACAGCAGTATGATGATCAAGGTACTTGGGGAATGACTTGGGTT 5915  
 QY 1847 g-----ggagagggggtctcagagctggaacttcaactcaactgactcaaaactggc 1894  
 DB 5914 GATGATACCGCAAGCACTCCAAAGACTTAAGACAGCGAATTAAGTGTGACCAATACAGAC 5855  
 QY 1895 tatattccatcccgagagctatcggtctttagtccctaaatagctttagaatgcatct 1954  
 DB 5854 TACCTTCGGAATCTGACGCGCAAGAGACCTTATGTTCTAATAGCCTTGGGGAATCTTT 5795  
 QY 1955 atagatatagctctcctcatatctttagagagactgcaaaacgaagggttgcagggagac 2014  
 DB 5794 TCAGACATCAAGCATTAAGGAGTGTCTATAGAGAAAGTGTGACCTTTGTTGTTCAAT 5735  
 QY 2015 cgtgcttttgggtggtggtatctactctctccaaaggtatgagtaagcaaaaaacaga 2074  
 DB 5734 CGAGGCTTCTGGGCTGCGGAGTCCCAATTTCTTAAGTAAAGATTAAGAAAGGGAAGAA 5675  
 QY 2075 cgggggttcgcaattgaggtggtggtatgcataggaagaaactacatactgttca 2134  
 DB 5674 CGCAATATCGCTCAATTAATCTGTGATATGCTATCGAGGTGACGCAAACTGTCTCT 5615







Db 1077 TCTATGACATAGATCTACTGCAAGATCAGAAATTAACGTGCAATATCTGGGCAATAG 1136  
 Qy 1249 cactactactatgatccataactacagatcatccacaacagttacagatgtcttaa 1308  
 Db 1137 CATCTTTTCTAGATCCGATTTACTCTAATACGGCTGGGAGATTCTACAGATCTTTAA 1196  
 Qy 1309 agttaatgagactccgagatctgcaactacaataacagagacatcatcttcacag 1368  
 Db 1197 TCTCAATTAAGGCTGATGAGGTATAGTACAGATATAGTGGTGCATGTTTCTGG 1256  
 Qy 1369 agaaaagttatcagagacagagcgccagatcttaaaactctactcgaagtactaa 1428  
 Db 1257 TGAAGAGCTCTCTGAGATGAGAACAAAGTTGAGCAACCTCATTCTACGCTGAGGA 1316  
 Qy 1429 gctctgaactcttcagagagatctactcttaaaacatgagatgagctcgaactca 1488  
 Db 1317 GCTGTACTCTAATGACAGGAATTTAGTACTTAACGTGGGTGACCTGATACGA 1376  
 Qy 1489 ggcattcactcaacagagatctcgtctcgaatgagacgtacgaactctagaacc 1548  
 Db 1377 AGGCTTTACTCAGACCGCGGTCTCTGTATATGATGCGGCAACGTTAAAGC 1436  
 Qy 1549 ---tgctgactagaccataaacaatttgcattacaatcagttctatagagctgc 1605  
 Db 1437 AGTACAGAGAGGTACTTTAACAGGTCTTTCATTCTCTAGACTCTTTAGCGAGGG 1496  
 Qy 1606 aagaagagcaaaataagaaaccaaagctacgtcaaaaactcgtactcttcagaaacc 1665  
 Db 1497 TAAGAAAGTTGTAATGTGCTTCTGCAAGTAATAAATGTAGCCCTTAGTGTCGAT 1556  
 Qy 1666 caattatgagacccgagcgacgctttatgaaactatgacttaagaactcctagtc 1725  
 Db 1557 TCTTCTTTGGATACCAAGAGGATCTTATGAAATCAAGATTGAGAAATCTTACAA 1616  
 Qy 1726 ctacgacatcttagagactcaaaagctctcgtgaactgaacagacccgcagtgactcaga 1785  
 Db 1617 CTTTCATTGTGACACTCTCTGCTGTGCTGATGCAACACACAGATGTTCCAGCGT 1676  
 Qy 1786 tccataatgagtgagaaatccatcaagcgctacagggagacttgggcccacttcttg 1845  
 Db 1677 TCTTACAGTAGAACACTCTACGACATGAGTATCAAGTACTTGGGGAATGACTGGGT 1736  
 Qy 1846 gg-----ggagaggggctctcagactgcaactcctcaacttgactcaaaactg 1893  
 Db 1737 TGTATGATCCGCAAGACACTCCAAAGACTAAGACGACATTGAGCAATACAG 1796  
 Qy 1894 ctatctccatacccgagcgatcagcgctctttagtccttaagttatggaatgactt 1953  
 Db 1797 CTACCTTCGGAATCTGACGCTCAAGACCTTATAGTCTTAATAGCCTTGGGATCTTT 1856  
 Qy 1954 tatagatatagctctctccatattcttatgagactgcaaacgaagggttgcaggaga 2013  
 Db 1857 TTACAGCATCCAAAGCATTCAGAGTGTATAGAGAAAGTGTGACTTTGTTGACAG 1916  
 Qy 2014 ccgtccttttgggtgctgagttatctactctccctaagagtagtacaacaaacag 2073  
 Db 1917 TCGAGGCTTCTGGCTGCGGAGTCCCAATTTCTTAAGTAAGTAAGAAAGGGGAAA 1976  
 Qy 2074 acgcgagtttcgcatlttagtggcggttatgcatagagagaaactacatactgttc 2133  
 Db 1977 ACGCAAAATACCGTCAATTAATCTGGTATATGATATCGAGGTGCGCCAAACTTGTTC 2036  
 Qy 2134 agataagattcttagtgcgcatcttgcagctcttggaaagatagagactacttctg 2193  
 Db 2037 TGAATACTTAATTAAGCTTTGCTTTGCAACTTTGTTAGCGAATAAGATTTCTTACT 2096  
 Qy 2194 agctaagaactcaaggtacagctcaagagagaaactctctacacaaagaaaccca 2253  
 Db 2097 CGCTAAATATCACTGATACCTATGACAGAGCTCTTATATCCACACATTAACAGATG 2156  
 Qy 2254 tatctctctcctgcaacactacgagcctgtctgtcttcttctctacagagattcc 2313  
 Db 2157 TAGTGGGTTCATAGGTTGTCTTATAGATAA---CTTCTGCTGCTTGGAGTCATTAACC 2213

Qy 2314 tgtctctttcaggaacacttagctacacccatacgaatacagatctgaaaccaa 2373  
 Db 2214 CCTCGTTTAAAGGAGGACCTGCTTATAGCCAGTCAATATGATGTAACAAAGTA 2273  
 Qy 2374 tacaacatactactgtttaaagaagctgggggaatgataagttctcgtttgaattcgg 2433  
 Db 2274 TATCGGTATCCTGAGGTGAAGGTTCTTGGGGATATAGCTTTTAAATGATGTTGGG 2333  
 Qy 2434 tgaagagctccgatttgcattagatgaaagtctctatttgagcagtaacatccctat 2493  
 Db 2334 ACCTCTCTCATTTTATCTGTAATACCTGCATTGTTTGAATACCTATGCTCATACAT 2393  
 Qy 2494 gaattgcagtttgcctatgcacacaggaaggttttaagaacagagagacagctcg 2553  
 Db 2394 CAATCGATATCGACTTATATGCTACAGACGCTTCTCGGAAAGTAACGAAAGAA 2453  
 Qy 2554 tgaattggaagtagccgctctgtgaattctgcttaccctacgagatccgattgataa 2613  
 Db 2454 ATCTTTTGTAGACACCAACCTCTCAATTTATCTTGGCTATAGGGGTGAAGTTGAGA 2513  
 Qy 2614 ggaatcagactccaagatgcaagctacaaactaactctgttctactgtgacttctgt 2673  
 Db 2514 GTTCTGTATTTAATGACTTTTCTTATGATCTGACTTATCTATGCTCTGATCTTAT 2573  
 Qy 2674 tctagtaaccccgactgacagacacactgcgaattagcagtgatgtcttgaacacct 2733  
 Db 2574 CCGCATATATCCCAATATGACATACAGCACTGTATACAGCGAGCTCTTGGGAACCTTA 2633  
 Qy 2734 cggtcgaatttggcaagaacagctttagctctcgttgcagaggaaacatttgccttaa 2793  
 Db 2634 TCCCATATACCTTAGACAGCAGCAGCCCTTCAAGTGGGTGAGCACTACACCTTCTC 2693  
 Qy 2794 ctcaattttagagcctttagcgaatttcttcttgaattgagtggtgactcgcgaata 2853  
 Db 2694 TCCCTATGTTGAAGGCTGCGGCACTTTGCTTTAAGTTCGTGATCCTGATCCTCAGATTTA 2753  
 Qy 2854 caatgtagacttagagacaaataacaaattcta 2886  
 Db 2754 TAATGTAGATCTTGGGGTAAGTTCCAATTTCA 2786

RESULT 12  
 CPN133034 17280 bp DNA 24-FEB-1999  
 ILOCUS  
 DEFINITION Chlamydia pneumoniae GCAT-protein family cluster A, genes omp12 (partial), omp11, omp10, omp5, omp4, omp3 and omp14 (partial) and ORF1, strain VR1310.  
 ACCESSION AJ133034  
 VERSION AJ133034.1 GI:4455881  
 KEYWORDS hypothetical protein; omp10 gene; omp11 gene; omp12 gene; omp13 gene; omp14 gene; omp5 gene; omp6 gene; ORF1; outer membrane protein; outer membrane protein 11; outer membrane protein 12; outer membrane protein 14; outer membrane protein 4; outer membrane protein 5.  
 SOURCE  
 ORGANISM Chlamydia pneumoniae.  
 Chlamydia pneumoniae.  
 Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 REFERENCE  
 1 (bases 1 to 17280)  
 Boesen, T.  
 Direct Submission  
 Submitted (21-JAN-1999) Boesen T., Department of Medical Microbiology and Immunology, University of Aarhus, The Bartholin Building, DK-8000 Aarhus, DENMARK  
 2 (bases 1 to 17280)  
 Hjerne, K., Boesen, T., Daugaard, L., Knudsen, K., Madsen, A., Christiansen, G., and Birkelund, S.  
 Chlamydia proteins containing the GCAT-repeat belong to a subfamily of auto-transferring pathogenicity factors  
 JOURNAL  
 FEATURES  
 source Location/Qualifiers  
 1..17280  
 /organism="Chlamydia pneumoniae"

```

/strain="VR1310"
/db_xref="taxon:83558"
1..841
/gene="omp12"
<1..841
/gene="omp12"
/note="putative"
/codon_start=1
/transl_table=11
/product="outer membrane protein 12"
/db_xref="GI:4455882"
/translation="EDNIRYRHNHSGGVALGTAATPAEDQLTEAFQGLFARPDNHTG
KMHGDTYGCASLYPHHTEGFDIANPLMKATRAPWYLSLSQILPSPDKSYLTD
NHKMTYITDNLISIKGWRNDACADGASLPFVLSVPLLEVEPVKQYITAAQOD
FYERHAGRAKNSSELIINVEITPVGTFEEDSKSEKGYIDLMTITDAIRRNPKQYIS
LIASDANMMAVIGINLARQGSVRAANHFOVNDHMEITFGQFAFEVRSRNYNTNLGSK
FCF"
885..968
/gene="omp11"
885..3677
/gene="omp11"
885..3677
/gene="omp11"
/note="putative"
/codon_start=1
/transl_table=11
/product="outer membrane protein 11"
/db_xref="GI:4455883"
/translation="MKTRPHKLLISLTLYTPILSLATYQADASLPTDFPDAGAST
FTPSITADAKNINVLSGNYITNDKSGALGACCFETTEDLFTFGKISFENYD
AGSNAGAASLTADKALITFGESNLSPFIAPETVYASGRSTLSAGALANTDNGTILP
SQVNSNEANNNGAITYTKTILSIGSTSTFTFNSNAKLGALYSSAGALNSISGNTGL
VFMNKEKTEGGALGFEASSITQNSLFEQSNTATDAGAGGALICEYTGEPPLTI
LSNSLSLFAENSVTGGACACAGDLASLPGTLESNNRCGNTAAGGAGALIASGS
LISANQDITFLGNLTLSAPTSTRAIYLGSSAKITMLRAAGQSLFYFDPLASN
ITGASDVLTINOPDNSPLDYSQTYVSGEKLSDEKAKADNFTSLIKQPLASGTL
ALKGVELDVNGFOTEGSTLLMOPTGKADTEALSTKLIVDLSALEGNKSVSET
AGANKTITLSPVFODSGNEFESHITNOAFTQPLVYFATAPADSIYIDALLSPV
OTPEHYGOGHMEATWADSTAKSGIMTWVTTGPNPERRAGVSDSLMAESTDIR
TLOOQNTSOANSIYOORGLASQTAAPFKHKSQNOAFRHSXGYVGGASDESDEN
IESVAFQGLFGKDKDLFIVENTSHNYLASLYLOHRAFLIGLPPMSGSGITDMKQPL
ILMOLASYTEKNDMDRTYSVPEAQSGTNNSGALTEGSLALYLPEKPEPFCQGF
FLKQAVYSROQNKPSGAPARAPDQGLYNCSTPGYRLEKTSPEDKNPELSAYI
GDYRKNDRSRTSLMVGASWISLCNNLARQAFIASAGSHITLSPHVELSGEATIELK
GSAHTINVGSLRTSF"
3827..6613
/gene="omp10"
3827..6613
/gene="omp10"
/note="putative"
/codon_start=1
/transl_table=11
/product="outer membrane protein"
/db_xref="GI:4455884"
/translation="MKSSLHWFLLISSIALPLSLNFAFAAWEINLGPNTSPSGPGT
YTPPOCTNADGTIYNLTGVSITNAGSPTALTSCEKETGNSLSPGSGXYOLONTI
DAGANCFPTNADANKLISFSGFSYLSLIOTNTATGCAIKSTGACISGNSYCTGO
NFSNDNGALOGSGSISLTPNITPAKNATOKGALYSTGCTINNTLISASFPETIR
AANNNGALYTBASSFTSKAISTFNNSYATASNGAIYSSSTSAKRPILTSDBGE
LNFIGNTLISGALITYTDNLVSSGGPFLKRNKSAIDTAAPLGGALIASGSLISA
LGGDITTEGRTVYKAGASSSTTRNSINIGNTNAIVLRASOGNTIYFDPTTISIT
AALSDALNKPDLAAGNPAIOGTIVSSEKLEAABADMLSTIQDPLTAAGQGLS
LKSQVTLAKFSQSPEGSTILMDAGTTLETAGITINVLVNDLSLKEKTKALAKQIS
ASQVTLGSLISLVDPGNYEDVSWNNNOVFSCLITLADPANIHLITDLAAPLEKN
PILHWGQGNMALSWOEDTATSKAATLTWTKGYNPNPERGRTIVANTLIGASVDVRS
IOQVATKVRQSOETRGIMCEGISNFFHKDSTIKNGKRHISAGYVVGATTTIASNL
ITAAFCQFLGKRDHPTINKNRASAYASLHQHLATLSPLSLAYLPGSSSEPVLEF
AOISATYCFKMTKTYTOAKPKGASWYNDGCALTELASLSPITALSHEGLHAPEYPIK
VEASVTHODSEKRENTLIVRSFSDGLIVSVPIGITERFSRNERASAYATVIYAD
VYRKNPDCTALLINNTSKTGTGTLNSRQAGIGRAGIFYAFSPLEVTLSLSEIRGS

```

```

sig_peptide
CDS
885..968
/gene="omp10"
complement(6954..8060)
/note="ORF1"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/db_xref="GI:4455885"
/translation="MSPQOPEQCHPDVYGSFLRPESLIRARSDEEGRITYEOWRY
EDAIRNLIRKQTEAGLIFPDGFRFYSQDPDPMFGSGCYRDPNDPRTGYLAD
KISVSKHPHEHFEVYKTFEGNNAKQOTIPSSQFPHMIRAPNLKMTKRFYPNQD
LIDDIYFTKQYIODLITAGCRNQLDDCACHRLDIDRAFMTGVDSDHRLQELIDQF
LWHLNLMKDRDEDLVSLHVCQKQYQAEFFSRATDSIEEPLFAKIDVDSHTYWL
DKRSGGAELAVYSGEKHVCGLISSHSCIEDRAVVSRTIEAASITPLERLSLSP
QCGFASCEDDHRETEEDMKRIAFVEIKAEIANG"
complement(8586..11372)
/gene="omp5"
complement(8586..11372)
/codon_start=1
/transl_table=11
/product="outer membrane protein 5"
/db_xref="GI:4455886"
/translation="MKSQPSWLVLSSTLACGTCSTYPAATAENIGSPDSFDGNNNG
TTPKNTTGTIDYTLTGDTTQNLGDSALATKCFQDTEGSLSPAGKYSPLNTS
SABGALSVYTDKNTSLNGPSSTIPFIAPSSVTPPSGAYKSGDGLFPPNKGITIF
KQDCEBNGGALISTKLSKSTGSLSEPEKSSAIGKKGALCATGYDITNNATP
LEBNILDEANGAINGSTGCTITGRTSLVSESVIRITKNGGALSQDADVTISGOS
VTEBGNANANGGALITAKKLTLAGSGGGGGSFSTNIVQGITNGGALSILLAGCSL
SAEGDITTEGNAIVATTPQTKRNSIDIGSTAKITNLRLISGHSITFYPDITANLA
DSYDITLNLKNAAGNSDYSGSIVFSEKSEDEARADNLSTLQKQYITLAAVLV
KRYGTLDTGFTQTAGSSVIMADAGTILKASTSEVYTLGSLIPYDLSGSKVYLAASA
AKSVVALSGPILILDNOGNAENHDLCKTQDFSVOLASLGAITTDVPAVPVATPT
HYGQGTGWGTVWDTASIPKTKATLAWNTGYLNPBERQGLVPNSLWGSFDTIA
IOGVIERSLATLCSDFGEMAGVANDYKDKRKGKRRKSGGVAIGGAQCSSEN
ISFAPCOLGSDSDCEFLVANKHNDYAGAFYLOHTTECGSITGCLDLKPGSKSRPLV
LEGGLAVSHNSDLKRTYAYDEVYGSWNNAFMMMLGASHSHPYELHCEPTYAPL
KLNTLYRQPSFESEKTEBRSFDSNENFSLPGVAFERFSQDNDESYDLISVDP
LIRNDPCTATLAVISGASMEYANNLARQALQVRAQSHVAFSMVEVLGQVFEYERS
SRINYNDIGKRFQF"
complement(11313..11372)
/gene="omp5"
11355..11586
/gene="omp4"
11355..14321
/gene="omp4"
11535..14321
/gene="omp4"
/codon_start=1
/transl_table=11
/product="outer membrane protein 4"
/db_xref="GI:4455887"
/translation="MKTSIPWIVSVLAASCHLOSANBELSPDPSFNGINDSGTF
TPKSATYTLGDFVEYEPKGNPLDSQCFKQTDPLFLGHSGLPGEFIDATTHA
GAASATANNKNTLFGSESLSPDSFSTYTTGQGLTASAGAVNENIRKIVANVS
TADGAKKASPLTGTSGDALFSNNSSTYTGAGLANTAGARLANTEIRKIVANVS
TSGGALIDBETSLSNKRKYIPEGKATTCGALIKGASGSESLISNKLITRAN
VARESGALIAKRLASSSGTPELRNNVSATPKGALSIDSGLSLSATGATTP
VRLTITLGTSTDPKRNALINGSKRFLKMAAHHITFIYDITSEGISVVALIN
GSABALPVGITLDSGETLADDELKADNLSFQVPSLSGKLLKQGSNIVYSGK
FSQAGSLIDMDSGTLLSTTAGSITITNLGINDSLGAKQVSLAKGASKNVIVSGK
LMLTIDEGNTYESHMSHOLESFLKTIIVADVDVNDISLIPPAEDPNVSEYEQG
QNNVNTDITATITKEATITWTKGFVPSPERKSAIVCNTLMGFEDIDSLQQLVEIG
ATGHEHQGFHWSSMTNPLHKTGDEKRGKFRHSGGVYIGSAHTEPKDILFAFACH
FAPKQDFLIHNNRSNTNGLTFKSHSNLQPCONTLRLGRARFESAEIEKFEIPLAL
DVYASFHSNDNREHTYSLSPESGSMNECJAGCIGLIDLPVLSPHRLPTEFTPM
KVEWYVYQNSPFESSSDQSGSISRLNLSIPGAKFVGQDGDSDSYTDLSGFEVSD
VYRNPDPSTATLIVSPDSMKIKINGMLSNQATLLAGSNVYVNSNCLETFHAMELRGS
SRNNNDVGYKRLF"

```

```

sig_peptide 14579..14644
/gene="omp13"
gene 14579..16123
/gene="omp13"
14579..16123
CDS
/gene="omp13"
/codon_start=1
/transl_table=11
/protein_id="CAB37073.1"
/db_xref="GI:4455888"
/translation="MTLRRLFCALFLPAAOVYIHESDYGNAINKSLKPK
ITCYPESTIIFLDVRIKVKHDOEDAGFIRSGNLFPMGRKCNFTHMLTEFG
AALSINRGDTLLSNFSYLAFLPGGALISLGVMLNSEEVTFCNYSWM
SGAIIYPTLLGSKASRSPSYNLGSRNRLVFRDVSQYGGAIHNLITLRPSCFE
NMHAYHDVNSGGAIAIAPGSGISIVKSGDLIFKGNIAODGNTIHNHIDGSCFE
KNLRVSEGVYFDPISHSKRTIDLVINAPGKRTYGTISFGGLCDHEVCAE
NLSTLIQDTYLAGTLLSDVTLQHSKQASSTLHNSPTTLLCSDAVQVH
IIEEDTNEVPYRIABDKDALYSLEKNAFPAVSVDPEPKAFNPLLELGG
SEDSLLEGTTLERTQTTENDAVRGFWLSNIEPSPSLDKRRIRIPTKYVLTWNP
EIRSTP"
sig_peptide 16494..16553
/gene="omp14"
gene 16494..17280
/gene="omp14"
16494..17280
CDS
16494..17280

Query Match 17.7% Score 530.2; DB 2; Length 17280;
Best Local Similarity 52.9% Pred. No. 1.e-103;
Matches 1415; Conservative 0; Mismatches 1183; Indels 75; Gaps 10;

QY 247 tgcctggaactaactactatattgaaggaaatgctacccctagaataatctccctgaacagg 306
DB 11217 TACTGGATAGACTATACCTGACAGAGATATACCTGCAAAACCTTGGGATTCGG- 11157
QY 307 cagacaatcacaaaagcgtctttaaacaacactaaggcgagattgactcttcacaggtaa 366
DB 11158 --CAGCTTTAAGCAAGGGTGTCTTCTGACACACGAACTTTAAAGCTTTGCCGTTAA 11101
QY 367 cgggaactctctatgttccaacagctgagtcagggactgacaggcgctgctgttaa 426
DB 11100 GGGGTACTACTCTTCTTTTAAATTAATTAAGTCTAGTGTGAAGCGCAGACACTTCTGT 11041
QY 427 cagagagtggtgatataatcaccacagttatagggcttctcgctacattatctgc 486
DB 11040 TACACTGATTAATAATGTGCTTAACAGATTTTCAGACTTCTTACGGGCCCC 10981
QY 487 gtctccggaagtcgataactacacagcgcaaggagccgttagctgcctcagggtagctt 546
DB 10980 ATCATGCGTATATCAACCCCTCAGAGAAAGTGCAATTAAATGTGAGGGGATC--T 10924
QY 547 gaagttgacaaaatgctagttgtctctcagcaaaaactttcaaggataatgagcg 606
DB 10923 TACATTGATTAACAAATGAATCTATTATTAAACAAGATTACTGTGAGGAAATGCGCG 10864
QY 607 tgcatacaccgcaaaaactttcattacaagggactacaaatgctcagctcgttttctga 666
DB 10863 AGCCATTTCACAGAAATCTTCTTGAAAAACAGCAGGGGATCGATTTCTTTTGAAGG 10804
QY 667 aaaba-----cctccctcaaaagaaagcgagaccattcagaactccagtcagcttac 717
DB 10803 GAATTAATCGAGCGCACAGGAAAGAAAGGCGGCTATTGTCTCACTGCTGATCTGTA 10744
QY 718 cattactggaacaaagggagagctcttcttcagacaatactctcagattctgagc 777
DB 10743 TATTACAAATAATACGGCTCTTACCTCTCTCGAACAAATATGCTCAAGCTGAGGTGG 10684
QY 778 tgcattttacagaagccctcggtgactattctcaataagtgtaaaagttccatttga 837
DB 10683 AGCATATAAATAGCAGAGAAATCTAATAATTACAGGAAATAGCTCTTGTGATTTTGTGA 10624
QY 838 caataaagtcacagagcgagctctcacaacaggggagataatgctcaggaagtgctatctg 897
DB 10623 AATAGTGTGAC--AGCGACCGCAGGAAATGAGGAGACTCTTCTCTGAGATGCGC---- 10569

```

```

QY 898 tgcattaaactagtagacagataactaagtcacccctcactcaggaatcagatgctactctt 957
DB 10570 -----ATGTTACCATATCTGGGAATCAGAGTGAATCTT 10537
QY 958 cagcaacaatatacgcacaacagcgaggagctatctatgtgaaaaagctcgaactgac 1017
DB 10536 CTCAGAGAAACCAAGCTGTACTATATGGCGGACATTATAGCTAAGAGCTTACACTGGC 10477
QY 1018 ttccggagagct-----taccctattcagtagaataatgtaagtaaggaagtcagctcc 1071
DB 10476 TTCCGGGGGGGGGGGGGTCTCTCTTTTCAACAATATAGTCCAAAGGTACCACTGACGG 10417
QY 1072 taagagtgagggcatagctatcgaagaatagtgagggaattgagtttccggcagatgag 1131
DB 10416 TAATGGGGGACATTCTTCTTACTGCGAGCTGAGAGTGTAGTCTTTACACAGAGCAGG 10357
QY 1132 tgcattgtctcttttagggataacagtcactctactactcc--tggagcaataagaag 1188
DB 10356 GACATTACCTTCAATGGGAAATGCCATTGTGCACTACACCAACAACATCAAAAAAGAA 10297
QY 1189 taatcgaacttaggaacgagtgcaagaatgacagcttgctgcgtcgcgtgtagagc 1248
DB 10296 TTCTATTGACATGAGATCTACTGCAAAAGATCAGAAATTTACGCAATATCTGGCATAAG 10237
QY 1249 catcactctatgatccataactaactaagatcatccacaacagttacagatgcttaaa 1308
DB 10236 CATCTTTTGTAGATGCGATCTACTGTATAGAGGTGCGGATTCATACGATCTTTAA 10177
QY 1309 agttaatgagactccggcagatctcgaactacaataacaggaacatccttcacagg 1368
DB 10176 TCTCAATTAAGCTGAGAGCAGATGATGATGATGATGATGATGATGATGATGATGATG 10117
QY 1369 agaaagttatcagagagacagagcgcaagttcctaaactctactctgaagtagctaa 1428
DB 10116 TGAAGAGCTCTCGAAGATGAAAGCAAAAGTTGCAACAACTTACTCTTACGCTGAAGCA 10057
QY 1429 gcctgtactcttctcaggaggtactatctttaaacaactgaggtgactctgacagctca 1488
DB 10056 GCCTGTACTCTTAACGACAGAAATTTAGTACTTAAACGAGGTGACCTGTGATACGAA 9997
QY 1489 ggcattcactcaacagcgagatctcgtctcgaatgagacgtggaactacttagaac 1548
DB 9996 AGGCTTACTCAACCGCGGGGTCTCTGTATTATGATGATGATGATGATGATGATGATGATG 9937
QY 1549 ---tgcataactagtagacataaacaattgctcattacatcagttcattagagcgtgc 1605
DB 9936 AAGTACAGAGAGCTCACTTTAAAGGTCTTCCATTCCTGTAGACTCTTTAGCGAGGG 9877
QY 1606 aaagaagggcaaaaatagaaaaccaaagctacgltcaaaaactctgaacttactgaaacct 1665
DB 9876 TAAGAAAGTTGTAATGTGCTCTCTGACGAAGTAAATAATGTAAGCCCTTAAGTGTGCGAT 9817
QY 1666 caatttatggaccggcagcgaggtttatgaaaaatcatalagtttaagaatcccaagtc 1725
DB 9816 TCTTCTTTTGGATTAACCAAGGAAATCTTATGAAATTCAGACTTGTGAAAACTCAAGA 9757
QY 1726 ctacagcatcttagagactcaaaagctctcgaactgtaaaagacgagcgtgctcaga 1785
DB 9756 CTTTTCATTGTGTGACACTCTCTCTCTGTGACTGCAACAACTACAGATTTCCACCGGT 9697
QY 1786 tccataatgggtgagaaatccatacaggtatcagggaaacttggggccaattgttg 1845
DB 9696 TCTTACAGTAGCAACTCTACAGCATATGAGGTATCAAGGATTTGGGGAATGACTTGGGT 9637
QY 1846 gg-----ggaacggggcctctcagactgtaaaccttcaacttgtaactaaactg 1893
DB 9636 TGAATGATACCGCAAGCACTCCAAAGACTAAGACAGGCAATTAAGCTTTGGAACCAATACAG 9577
QY 1894 ctatattcctaatcccgagcgtagtcgctctttagcctccttaagctttagatgcat 1953
DB 9576 CTACCTTCGGAATCCGAGAGGTCAAGAGACTTATAGTTCTAATAGCCTTTGGGATCTTT 9517

```

[illegible]

FEATURES	source
1. 534	/organism="Chlamydomophila pneumoniae"
/strain="j138"	
/db_xref="taxon:83558"	
/note="synonym:Chlamydia pneumoniae"	
1. 534	
/gene="pmp-3.1"	
1. 534	
/gene="pmp-3.1"	
/note="Containing frame shift"	
/codon_start=1	
/transl_table=11	
/product="polymorphic outer membrane protein G family"	
/protein_id="BA08947.1"	
/db_xref="gi:6172315"	
/translation="MKSSFPEKVESTRAIFPLSMIAETVLDSSASFDGKNGFSVRAESQEDACTTTFKGNVILENIPGTGTAITKSCFENNTGDLTFITNGNSLTFQVADGTVAGAAANSVVDKSTFTFGFSSLSFIASPGSSITTGKAVSCSTGSLTKMSVCSAATKTFRIANVSLSPDLF"	
BASE COUNT	147 a 105 c 117 g 165 t
ORIGIN	
Query Match	17.4%; Score 523; DB 1; Length 534;
Best Local Similarity	99.8%; Pred. No. 5e-102;
Matches 534; Conservative	0; Mismatches 0; Indels 1; Gaps 1;
101 atgaagctctctccccaagttgtattcttcacatttgcatttcccttgcctatg	160
1 ATGAAGCTCTCTTCCCAAGTTGTATTCTTCATATTTGCTATTTTCCCTTGTCTATG	60
161 attgctaccgagacagtttgatcgaatgcgaattcgaatggaataaataatggtat	220
61 ATTGCTACCGAGACAGTTTGGATCAAGTCGAGTTTCGATGGGAATAAATGCTAT	120
221 tttaagtcgtgagagcaggaagatgctggaactactcctcatttaaggaatgct	280
121 TTTAGTTCGTGAGAGTCAGGAAGATGCTGGAATCACTATTTAAAGGAATGTC	180
281 actctagaataatctctcggaacaggcacaacatcaacaaagctgtttaacaaact	340
181 ACTCTAGAAATATTTCCGGAACAGGCACACAAATCAAAAACCTGTTTAAACAACCT	240
341 aaggcgagtttgacttcacaggtaacagggaactctctatgttccaacagtgatgca	400
241 AAGGCGAGTTTGACTTTCACAGGTAAAGGGAACCTCTATGTGTCCAAACGCTGATGCA	300
401 ggaactgaagcagagggatgctggttaacagcagcgttggtagaataatccacgttata	460
301 GGGAAGTGAAGCAGGGGCTGCTGTAAACAGCAGCGTGGTACGTAATATATCACAGCTTATA	360
461 gggattctcttcgcatcatttattgctgctcctcggaagttcgataaactaccggcaagga	520

Db 361 GGGTTTTCCTCGTATCTTTATTCGCTCCTCGAAGTTCGATACCTACCGGCAAGA 420  
 QY 521 gccgttagtgcctctacgggttagcttgagttgacaaaatgctcagttcctcagc 580  
 Db 421 GCGGTAGCTGCTCAAGGGGTAGCTGAGTTGAC-AAAAGTCAAGTTGCTCTCAGC 479  
 QY 581 aaaaatttcaacgataatgctgctcctacacgcaaaactttcatiaa 635  
 Db 480 AAAAATTTCACAGGATATGCGGTGCTATACCGCAAAAACCTTTTCATTAA 534  
 RESULT 14  
 A81837 2757 bp DNA PAT 21-JAN-2000  
 LOCUS Sequence 11 from Patent WO9858953.  
 DEFINITION A81837  
 ACCESSION A81837  
 VERSION A81837.1 GI:6731869  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 unclassified.  
 unclassified.  
 unclassified.  
 1 (bases 1 to 2757)  
 REFERENCE  
 AUTHORS Madsen, A. and Birkelund, S.  
 TITLE NOVEL SURFACE EXPOSED PROTEINS FROM CHLAMYDIA PNEUMONIAE  
 JOURNAL Patent: WO 9858953-A 30-DEC-1998;  
 MAUSEN ANNA SOFIE (DK); BIRKELUND SVEND (DK)  
 FEATURES  
 source 1..2757  
 Location/Qualifiers  
 /organism="unclassified"  
 /db\_xref="taxon:32644"  
 BASE COUNT 770 a 564 c 642 g 781 t  
 ORIGIN

Query Match 17.4%; Score 522.2; DB 5; Length 2757;  
 Best Local Similarity 53.4%; Pred. No. 6.4e-102;  
 Matches 124; Conservative 0; Mismatches 1153; Indels 90; Gaps 12;

QY 242 gaagatgcggaactacactatcttaaggaatgctacactctagaatatctcctga 301  
 Db 160 GATGCTAGTGGCAGCAGCCTATATTCGATGGGAGTCTCGATA--AGCCAAGAGG 216  
 QY 302 aagggacacgacatcaacaaagctgtttaaacaacactaaggcgatttgactaca 361  
 Db 217 AAACAACACAGCTTAACCAAGTTGTTTCTAACACATGCAAGGAATCTTACTCTTA 276  
 QY 362 ggttaagggaactctctatctgttccaaagcgtgagtcaggagctgtacaaaggctgtc 421  
 Db 277 GGGACGAGATTTCCTCTTATTTGACATATATTTCTGCTACTGTTCCAGGTGTTGT 336  
 QY 422 gtttaacagcagcgtggttagataatctaccacgctttatagcgttttctcgcctatctt 481  
 Db 337 GTTAGCAATACAGACAGCTCTGCGATTACGAATTCAGGATTTTCACTCTTCGGAG 396  
 QY 482 attgctctccctcggaagcttcgataactaccgcaaaagcgttagcgtcctcagcgt 541  
 Db 397 CTTCGACGCTCT-----AGGACACAGAGTAAAGGAGCCATTA--AAATTACCGAT 444  
 QY 542 agcttgagttgacaaaatgtcagttgtccttcagcaaaaactttcaacgataat 601  
 Db 445 GGTCTGCTTTGAGAGTATAGGAACTTTGACCAAAATGAAATGCTCTTGTGTAAT 504  
 QY 602 ggcggtgtatataccgcaaaaactttcatatacagggagtaacatgtcagctgttt 661  
 Db 505 GGGGGGACCATTAATACGAAGCTTGTCTTACGAGGAGGAGGCTTGTAGAGTTC 564  
 QY 662 tctgaaaataactcctcaaaagcgtgagcattcagaattccgattcccttaacct 721  
 Db 565 CTTCGCAATAGCTCGTCGACACAGGGGAGGAGCATCTAAGCTTGTGCTACCTGTGAT 624  
 QY 722 actggaacaaaggaagtcctctttctgcaaatctcttcgagattcgtgagctgca 781

Db 625 TCTGAATGCAGGATCTTGAGCTTCGGAACAACAGTGGCAACAATCAGAGCGCG 684  
 QY 782 attttacagaagcctcgttgtaactatcttaataatgctaaagttcccttaatgacaat 841  
 Db 685 AACTCTGCTGAAGGGAGACCTTGATCTCCATAATCCAAAATATCTTTTGATGGCTGC 744  
 QY 842 aaggtcacagggcagctccctcaacagcggggatatgtcagggagcgtatcctgtc 901  
 Db 745 AAGCAACTACAAAGCGGAGCTTATGATTGTAACAAGCAGGG----- 790  
 QY 902 tataaactgtacagataaaggtcaccctcactcagaaatcagaigttaactctcagc 961  
 Db 791 -----CGAACCCAGACCTTATCTTGACTCTTTCAGGAAATGAGAGCTGATTTCTG 843  
 QY 962 aacaatactgcacaaacgcyggagagctatctatgtgaaaagctcgtacgtctcc 1021  
 Db 844 AATAACACAGCAGGAATAGTGAAGTGCATTTTACCAAAAATTTGTTATCTCA 903  
 QY 1022 ggaagacttaccctatctcagtagaatagtgtcaatgtgaggtacagctccaaaggtga 1081  
 Db 904 GAGCAGAGAGAGTGTATTTTCTAACACAAGCTCGGATGCTACTCTTAAGAGAGG 963  
 QY 1082 gccatagctatcgaagatagtgaggaaattgagttatccgcgcgaatgtgtacattgc 1141  
 Db 964 GCATTTGGGATCTGATTTCTGAGAGATAGCATTTCTGAGATCTCGCAATATCAT 1023  
 QY 1142 tttttaggaatacagctactctac-----tactcttgagcgaatagaatgtatc 1195  
 Db 1024 TTCGAGGGCAATACACAGCAGCCTACAGAGTCTGAGAGTGCACCAAAATCTTATA 1083  
 QY 1196 gacttaggaacagtgacaaagatgacagcttgctgcgtcgtgtgtagagccactac 1255  
 Db 1084 GATCTTGATGCAATGCAAAATTTTAATCTCCAGAGGACTCGGGGAATAATGATAT 1143  
 QY 1256 tctctatgccataactacagagatcaccacaacagttacagatgtcttaaaagtta 1315  
 Db 1144 TTCTATGATCCATACAGA-----GCTCAGAGCTACTGATGAACCTCTTTGAT 1194  
 QY 1316 ggaactccgcagatctgcactacaatatacagaagaaactatctacaggggaaag 1375  
 Db 1195 AAGCTGACGCGAGATCTGGAATACCTATGAAAGCTTCAATGTTTCTGAGAGAA 1254  
 QY 1376 ttatcagaagacagagccgcagatcttaaaaacttactcgaagctacacagcttga 1435  
 Db 1255 CTCTCAGAGAAGAACTTAAGAACTGACATCTGAAGTGTACATTACACAGGCTTA 1314  
 QY 1436 actcttcaagaggtactctatctttaaacaatgaggtgactcgcagactcagcattc 1495  
 Db 1315 GAGCTTGTGCGAGGTGCTTATGATTAAGATGAGAGTGTGTTGCAAAATCTATA 1374  
 QY 1496 actcaacagcagatctcctcgtctcgaatgagcaggaactacactga---actcgt 1552  
 Db 1375 ACGCAGTCCAGAGATCCAAAGTCTTATGATGAGAGGACTACTTTTGAGGCAAGCCT 1434  
 QY 1553 gatactagaccataaacaatttggcataacatcagttctatatacagcgtgtcaagaag 1612  
 Db 1435 GAGGGGTCACCTCAAGGCTAGCCATTAATAGATTTCTTGAATGGGCAAAATTA 1494  
 QY 1613 gcaaaaatagaacccaagcctacgtcacaataatctgacttatctcgtgaacatactta 1672  
 Db 1495 GCTATCATTTAAGGCGACGCGCAAGTAAGATGTGCTTATCAGGGCTCATATGCTT 1554  
 QY 1673 ttgacccgacggcagctttaaataatagatttaagaataatccctcagctcagaac 1732  
 Db 1555 GTAGATGCTCAGGGGAATATATGAGCATATATATCAATCAACAGCAGTCTTTCT 1614  
 QY 1733 attcagaagctcaaaagctctggaactgtaaacagcagcgagtgactcagatcctata 1792  
 Db 1615 TTAATAGAGCTTTCTGCAAGAGAAAGATGACTCATACAGATATCCCGATACCCAAAT 1674  
 QY 1793 atggttgagaattccattagagctatcaggaacttggtggccaattgtttggggaca 1852  
 Db 1675 CTAATACTACGATCACTATGAGTATCAGGAAGT---GGAATTAATGTTGAGGTGCAC 1731

QY 1853 gggagctct-----acgactgcaaccttcaactgcaataaactggctatctcct 1903  
 Db 1732 GATGCAACTCAAAAACAAAAATGCTACCTTAACCTTAAGTGAACAGAGATCAAGCCG 1791  
 QY 1904 aatcccgaggtatcggctctttagctccctaataagctttagatgataatgataat 1963  
 Db 1792 AATCCAGAAAGCTCAGGAGACTTTGGTTCCATATAGCCGTGGGGTCTTTTGTGCAATGTC 1851  
 QY 1964 agctctccattatcttattgagactgcaacgaaagggttgcaggagagacgctgctt 2023  
 Db 1852 CGGTCAATTGAGACCCATGACGACCGGAGACAAGTTGCTTACTGTCGCAACAATTTG 1911  
 QY 2024 tgggtgctgacttacttactctctccataagataagatacaaaaacgacgctgctt 2083  
 Db 1912 TGGGTATACAGAAATCCGCGACCTTTTGCATGAAGATCAAGAAAGAAACCAAGCTATAT 1971  
 QY 2084 cgcacattgagtgcggtatgcatagagagaaactacatactgttcagataagatt 2143  
 Db 1972 CGTCAATTGAGCCGGGCTATGCAATAGAGAGATCTTACAGGCTTCTGAAAATTTTC 2031  
 QY 2144 cttagtgcgacttctttagactcttggaaagatagagactacttggtaagataat 2203  
 Db 2032 TTTAATTTTGTCTTTGTAGCTTTTGGCTACGACAAAGACCATCTGTGCTAAGAAC 2091  
 QY 2204 caaggtacagctcagagaaactcctctatccacagcaacgaaacctatctctct 2263  
 Db 2092 CATACCATGATATATCAGAGGGCAATGATTAACGACACCTCGGAGACTCTAAGACCTTC 2151  
 QY 2264 ccttgcaaaactacgacctgtctgtcttcttctcacaagagatctcgtctctt 2323  
 Db 2152 GCTAAGA-----TTTGTGAGGAAATCTGACTCCCTACCTTTGTCTTC 2196  
 QY 2324 tcaagaaacctagctacacccatacgcgataacgatacgtctgaaacccaatatacaat 2383  
 Db 2197 AATGCTCGGTTGCTTATGCGCATACGACATTAACATGACCAACAAGTACACCTGGCTAT 2256  
 QY 2384 cctactgtttaaagaaagctggggaagatagattcgcctttagaattcgttgcgaaagct 2443  
 Db 2257 TCTCTGTTAAGGGAAGCTGGGGAATGATGCTCGTATAGAAATGAGGAGACTATC 2316  
 QY 2444 ccga---tttgcttagaagaaagtgtctctatttggcagctacatgccttcaatg 2500  
 Db 2317 CCGAGTAGTGTCTGACGACGCTGCTGGGTGGATACCCACACGCAATTTCTAACCCTA 2376  
 QY 2501 cagttgtctatgcaacacgaaagttttaaagacagagaaagctcgtgaattt 2560  
 Db 2377 GAGATGATCTATGACATCAAAATGACTTTAAGGAAAACGCAAGAAAGCCGTTCTTC 2436  
 QY 2561 ggaagtagcgcgtcttgaactctgacctactacgtcggaatcgaaattgtaagaaatca 2620  
 Db 2437 CAAGTGAAGACCTTCTTAATCTAGCGGTCTCTAGGGAATAAATTTAGAA-----A 2490  
 QY 2621 gactgcaagatgcaacgatacaactctgctttagtactagtgtgacttctgttcgagt 2680  
 Db 2491 TTTCCGATAGTCTAGCTATGATCTCTCCATACCTTAACGTTCCCGATGTGATTCGTAAT 2350  
 QY 2681 aacccgacgtgtaagaaacacgctgcaatagcgggtgacttcttggaaacacttcgtacg 2740  
 Db 2551 GATCAAGGCTGACGACAACTCTATGTTCTGGGGATTTCTGGTGCACATGGGTACA 2610  
 QY 2741 aatttggcaagacaagctttagtctctcgtgcgaggaacacatttggcttaactcaat 2800  
 Db 2611 AGCTTGCTAGACAAAGCTCTTCTGTACGTGCTGGGAATCATATGCTTTGCTCAAAAC 2670  
 QY 2801 ttggaacctttagcacaatttctttagaattgctgtgctcactcgcacataatgata 2860  
 Db 2671 TTGGAATTTTCACTAGTCTGTTGAAGTGAAGTGCAGAGTCTTCTGTAAGCTATGCTATC 2730  
 QY 2861 gacttagagcaaaataacacattctaa 2887  
 Db 2731 GATCTTGGAAGAAATTCGGAATTTAA 2757

RESULT 15  
 LOCUS AE002193/c  
 DEFINITION Chlamydia pneumoniae AR39, section 26 of 94 of the complete genome.  
 ACCESSION AE002193  
 VERSION AE002193.2  
 KEYWORDS GI:8163406  
 SOURCE  
 ORGANISM Chlamydia pneumoniae AR39.  
 Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 REFERENCE  
 1 (bases 1 to 10026)  
 Read, T.D., Brunham, R.C., Shen, C., Gill, S.R., Heidelberg, J.F., White, O., Hickey, E.K., Peterson, J., Umayam, L.A., Utterback, T., Berry, K., Bass, S., Linher, K., Weidman, J., Khouri, H., Craven, B., Bowman, C., Dodson, R., Gwinn, M., Nelson, W., DeBoy, R., Kolonay, J., McClarty, G., Salzberg, S.L., Eisen, J. and Fraser, C.M.  
 Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39  
 Nucleic Acids Res. 28 (6), 1397-1406 (2000)  
 2 (bases 1 to 10026)  
 Read, T.D., Brunham, R.C., Shen, C., Gill, S.R., Heidelberg, J.F., White, O., Hickey, E.K., Peterson, J., Umayam, L.A., Utterback, T., Berry, K., Bass, S., Linher, K., Weidman, J., Khouri, H., Craven, B., Bowman, C., Dodson, R., Gwinn, M., Nelson, W., DeBoy, R., Kolonay, J., McClarty, G., Salzberg, S.L., Eisen, J. and Fraser, C.M.  
 Direct Submission  
 Submitted (01-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA  
 On Jun 1, 2000 this sequence version replaced gi:7189234.  
 LOCATION/Qualifiers  
 1..10026  
 /organism="Chlamydia pneumoniae AR39"  
 /strain="AR39"  
 /db\_xref="taxon:115711"  
 /note="synonym: Chlamydia pneumoniae AR39"  
 complement(105..2897)  
 /gene="CP0307"  
 complement(105..2897)  
 /gene="CP0307"  
 /note="similar to GP:4376729; identified by sequence similarity; putative"  
 /codon\_start=1  
 /transl\_table=1  
 /product="polymorphic membrane protein G family"  
 /protein\_id="AF38164.1"  
 /db\_xref="GI:7189235"  
 /translation="MKTPHKLISLTVLPILISITVYGAASLSPTDFDAGSSTFTPKSTADNGTNYLSGNYINDAGGTALGCCPEFTGDLTFPGKGSFSPNVDAAGSNAASTADAKALTFGFSNLIAFGTVAAGSTLSAGLNLDTGTFESQVNSNANNGGAIITAKTILISGNTSIFATFQPLAKLGAALYSASISGNTOLFVNKKGTGGGALGEPASSITONSILFPGMDADPAAGKGAITCEKGPNTLISGKSLTFAENSVTGGACANGLDLSAGPPLFSNRGNRAAGKGAIAIADSGSLSLANOGLITFGLNTLSTAPSTNATILYSSAKITLURAQQSITFYPIASN TTGASVILITNOPSNRPDIYSGTTFVSGKLSADEKADNFTSLIKOPLAASGTLAKNVAELDNGFLOTGSLTLMOPGRIKLADAEALSLTLVLDLALIDENKSVETAGAKTITLISPLVFDOSGNTFESHINAFQPLVFAAIAADIDYIDALTSFVQPEPHYGOHWEATWADSTAKSGMTWVTGYNPNERRASVYDPSIMAFDRI TLQOIMTSQANSIYIOGRGLWASGTANFHKDKGTQAOPRHKSGYIVGSADESEN IFSVAFQOLFGRKDKLEIVENTSHNYLASIXLOHRAFGLGPRPSGSLTDMKDIPLIFKFOAYSRQNFKESGAEARAFDDGDLVNCISIPVGRLEKISEDEKNNFEISLAYI GSAHIVVDCGLVYSF"  
 /gene="CP0308"  
 complement(2941..5751)  
 /gene="CP0308"  
 complement(2841..5751)  
 /gene="CP0308"  
 /note="similar to GP:4376728; identified by sequence similarity; putative"  
 /codon\_start=1





QY 1498 tcaaacagcagattctcgtctcgaatggaagtaagaaactactct--agaacctgtga 1554  
 Db 1506 ACAGACTGAAGGCTCTACACTCTCATGCAACAGGAACAACAAAGCGATATCGA 1447  
 QY 1555 tactagccaccaaaccaatttggatcaacatcagttctatagccgtgtcaaaagc 1614  
 Db 1446 AGCTATCAGTCTTACCAAACTGTGTGATTTCTGCTTAGGGGAAATTAAGTGT 1387  
 QY 1615 aaaaatagaaccaaagctacgtcaaaaatctgaacttattctgaaacatcaacttatt 1674  
 Db 1386 GTCCATTGAACAGCAGAGGCCAACAAACTTAACCTTAACCTCTCTGTTTTCGA 1327  
 QY 1675 ggaaccgagcgagcagttttttgaaatcatagtttaagaatctcagtcctacagat 1734  
 Db 1326 AGATGATGCGGCAATTTTATGAAAGCCATACGATAACCAAGCCTTCAAGCCCTTT 1267  
 QY 1735 cttaagctcaaaagctcttgaactgt-----taacaagcaaccgagtgactccaga 1785  
 Db 1286 GGTGATATTCACTGCTGCTACGCTAGCGATTTTATTCGATGCGCTTCACATTC 1207  
 QY 1786 tccataatgggtgaagaatccatcactacgctatcaggaacttggggcccaattgttg 1845  
 Db 1206 TCCAGTACAAATCCAGAACCTCATTAAGGGTATCAGGGGACATTTGGAGCCACTTGGGC 1147  
 QY 1846 ggggacaggggctctacagctgcacacttcaactgcgaactaaactgactatctctaa 1905  
 Db 1146 AGACACATCACTGCAAAATCAGGACATATGCTTGGGTAATACGCGGCTACAACTTAA 1087  
 QY 1906 tcccgagcgtatcggtctttagtccctaataagcttatgaaatgcatlataagattag 1965  
 Db 1086 TCCTAGAGGCTAGAGCTTCCGATTTCCGATTCATTAATGGGCATCCTTACTGACATTCG 1027  
 QY 1966 ctctctcattatctttagggagactgcaagaaaggttgcagggagacggtcttttg 2025  
 Db 1026 CACTCTACAGAGATCATGACATCTCAGCGAATATGATCTATCAGCAACGAGACTCTG 967  
 QY 2026 gtgtcgtgattatctactctctcacaagaagtagatgaaacacagacgaggttcg 2085  
 Db 966 GGCACTACAGGAACCTCGAATTTCTTCCATTAAGATTAATCAGGAATCAACCAAGATTCCG 907  
 QY 2086 ccatttgaagtcggttatgtcatagaaggaacatactatgttcagataagattct 2145  
 Db 906 ACATTAAGCTACGGCTATATGTTGAGGAGAGTGAAGATTTTCTGAAAAATATCTT 847  
 QY 2146 tagtgcgtcatttgcagctctcttggaagatagatagactactttagtaagatca 2205  
 Db 846 CAGTGTAGCTTTCGCCAGCTCTCGGTAAAGATTAAGACCTGTTTATAGTTGAAAAATAC 787  
 QY 2206 aggtacagctcagcgaagaaactctataaccagcaacgaacatactctctcc 2265  
 Db 786 CTCTCATTAATATTAGCTGTGCT---ATACCTGCAACATCGAGCATTCCTAGAGGACT 730  
 QY 2266 ttgcaaacctacgacctgttcgttcttattgttccctacagagattccgttctcttc 2325  
 Db 729 TCCCATGCCCTCATTTGGAAGTATCACGCAATGCTGAAAGATATTCCTCATTTTGA 670  
 QY 2326 aggaacacttagctacacccatacgaataacgaatctgaaacaaagtaatacaacatacc 2385  
 Db 669 TGCCCACTAAGCTACAGCTACACATAAAATGATATGATACCTCGCTATATCTCTATCC 610  
 QY 2386 tactgttaaggaagctggggaatgtagtttcgctttagaattcgttgaaga--gc 2442  
 Db 609 TGAAGCTCAAGGCTCTTGACCAATAACTCTGGGCTCTAGAGCTCGAGAGATCTGGG 550  
 QY 2443 tccgattttagatgaagaagtgctctattttagcagtagatcccttcatgaattgca 2502  
 Db 549 TCTATATCTCCCTAAAGACACACCTGTTCCAGGATATTTCCCTTTAAAGTTCCA 490  
 QY 2503 gtttgcctatgcacatcagaagaagttttaagaacaggaagaagctcgtgaatttg 2562  
 Db 489 GGCAGTCTACAGCCGCCAACAAACTTAAAGAGAGTGGGCTGAAGCCCGTCTTTTGA 430

QY 2563 aagtagccgtctgtgaatcttgccttacctatcggagatccgatttgataaggaatcaga 2622  
 Db 429 TGATGAGAGACTAGTGAACGCTCTATCCCTGTGCGCATTCGGTTGAAAAAATCTCCGA 370  
 QY 2623 ctgcgaagatgcaacgctacaaactbaactctgttatactgtggaatcttgcgtagtaa 2682  
 Db 369 AGATGAAAAAATTAATTCGAGATTTCTTAGCCCTACATGTGATGATGATGTAATAA 310  
 QY 2683 ccccgactgtacgacaaactgcaatttagcgttgatttcttggaaaaacttcggtacgaa 2742  
 Db 309 TCCCGCTTCCGCTACTTCTTAATGCTCACTGAGAGCCCTTGAGACTTCGCTATGTAATA 250  
 QY 2743 ttggcaagaacagctttagtctctcgtgcaggaagaacatttgccttaactcaaat 2802  
 Db 249 CCTGCGACGACAAAGCCTTCTTAGCAAGTGTGGAACCAATCTGACTCTCCCTCATGT 190  
 QY 2803 tgaagccttagccaatttcttgaattgcgttgatcgttgatcctcgcaattacaatgtaga 2862  
 Db 189 AGAATCTCTGGGGAAGCTGCTTATAGACTCTGCTGCTCAGCACACATCTAATAATGTA 130  
 QY 2863 cttagagcaaataccgaattctaatgct 2892  
 Db 129 TTGTGGCTAAGATCACTCATTTAGTTCT 100

Search completed: November 20, 2000, 09:30:28  
 Job time: 38600 sec



Tue Nov 21 11:10:18 2000

us-09-428-122-1.rge

Page 33



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 23:29:48 ; Search time 133.98 Seconds  
(without alignments)  
8411.614 Million cell updates/sec

Title: US-09-428-122-1

Perfect score: 3000  
Sequence: 1 gcgcctaccctagtagaggt.....tgcttgctaaacacttc 3000

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N.Geneseq\_36:\*

- 1: /SID6/gcgdata/geneseq/geneseq/NA1980.DAT:\*
- 2: /SID6/gcgdata/geneseq/geneseq/NA1981.DAT:\*
- 3: /SID6/gcgdata/geneseq/geneseq/NA1982.DAT:\*
- 4: /SID6/gcgdata/geneseq/geneseq/NA1983.DAT:\*
- 5: /SID6/gcgdata/geneseq/geneseq/NA1984.DAT:\*
- 6: /SID6/gcgdata/geneseq/geneseq/NA1985.DAT:\*
- 7: /SID6/gcgdata/geneseq/geneseq/NA1986.DAT:\*
- 8: /SID6/gcgdata/geneseq/geneseq/NA1987.DAT:\*
- 9: /SID6/gcgdata/geneseq/geneseq/NA1988.DAT:\*
- 10: /SID6/gcgdata/geneseq/geneseq/NA1989.DAT:\*
- 11: /SID6/gcgdata/geneseq/geneseq/NA1990.DAT:\*
- 12: /SID6/gcgdata/geneseq/geneseq/NA1991.DAT:\*
- 13: /SID6/gcgdata/geneseq/geneseq/NA1992.DAT:\*
- 14: /SID6/gcgdata/geneseq/geneseq/NA1993.DAT:\*
- 15: /SID6/gcgdata/geneseq/geneseq/NA1994.DAT:\*
- 16: /SID6/gcgdata/geneseq/geneseq/NA1995.DAT:\*
- 17: /SID6/gcgdata/geneseq/geneseq/NA1996.DAT:\*
- 18: /SID6/gcgdata/geneseq/geneseq/NA1997.DAT:\*
- 19: /SID6/gcgdata/geneseq/geneseq/NA1998.DAT:\*
- 20: /SID6/gcgdata/geneseq/geneseq/NA1999.DAT:\*
- 21: /SID6/gcgdata/geneseq/geneseq/NA2000.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3000	100.0	3000	A27021	Chlamydia pneumoni
2	2975.4	99.2	1230025	X91990	Nucleotide sequenc
3	2775.8	92.5	2787	X06820	Chlamydia pneumoni
4	530.2	17.7	2815	X06817	Chlamydia pneumoni
5	524.6	17.5	2957	21	DNA encoding the C
6	522.2	17.4	2757	20	Chlamydia pneumoni
7	515	17.2	2787	20	Chlamydia pneumoni
8	514	17.1	2793	20	Chlamydia pneumoni
9	506	16.9	3000	20	Chlamydia pneumoni
10	465.8	15.5	3200	20	Chlamydia pneumoni
11	389	13.0	3050	21	DNA encoding the C
12	378.4	12.6	2838	20	Chlamydia pneumoni

13	340.2	11.3	2526	20	X06819	Chlamydia pneumoni
14	273.2	9.1	3052	20	X06818	Chlamydia pneumoni
15	243.4	8.1	3150	21	A27342	Chlamydia pompi
16	209.2	7.0	1830	20	X06830	Chlamydia pneumoni
17	197.6	6.6	840	20	X06824	Chlamydia pneumoni
18	140.6	4.7	3354	20	X60556	Seq ID No: 23 of W
19	133.8	4.5	3324	20	X60557	Seq ID No: 24 of W
20	132.2	4.4	4435	20	X60539	C. trachomatis LGV
21	130.6	4.4	1444	20	X60547	C. trachomatis HMW
22	104	3.5	1200	20	X06829	Chlamydia pneumoni
23	80.8	2.7	1545	20	X06825	Chlamydia pneumoni
24	50.4	1.7	1511	20	X60546	C. trachomatis HMW
25	50.4	1.7	1515	20	X60555	Seq ID No: 22 of W
26	46	1.5	787	20	X06826	Chlamydia pneumoni
27	43.4	1.4	1038602	20	201425	Complete genome se
28	42.4	1.4	3975	9	N81157	Malaria-specific 9
29	42.4	1.4	3975	13	Q2299	SERP gene. Plasm
30	42.4	1.4	6124	11	Q03568	Sequence encoding
31	41	1.4	1497	21	256377	Escherichia coli f
32	41	1.4	1689	8	N70495	Sequence of Escher
33	41	1.4	2108	10	N91116	Flagellin gene of
34	41	1.4	2109	8	N70985	E. coli flagellin 9
35	39.6	1.3	580073	18	T58840	Mycoplasma genital
36	37.6	1.3	2963	18	T93775	PSD-93 coding sequ
37	36.8	1.2	1269	19	V12774	Metal-regulated tr
38	36.4	1.2	1598	20	X85043	Human secreted pro
39	36.2	1.2	1300	21	261504	DNA encoding the C
40	36.2	1.2	3050	21	261586	Nucleotide sequenc
41	36	1.2	2610	13	Q32556	LEMM1 8 CDNA. Lyc
42	35.8	1.2	1951	16	Q94449	Bacterial transfer
43	35.8	1.2	1951	18	T49507	Transferin recept
44	35.8	1.2	1951	19	V21443	H. influenzae stra
45	35.8	1.2	1951	21	289171	H. influenzae non-

#### ALIGNMENTS

RESULT 1	
ID A27021	standard; DNA: 3000 BP.
AC A27021:	
DT 11-AUG-2000	(first entry)
XX	
DE Chlamydia pneumoniae 98kD putative outer membrane protein gene.	
XX	
KM Chlamydia; antigen; vaccine; infection; outer membrane protein; ds.	
XX	
OS Chlamydia pneumoniae.	
XX	
FH Key	Location/Qualifiers
FT CDS	101..2887
FT	/tag- a
FT	/product- "98kDa putative outer membrane protein"
XX	
PN WO200026237-A2.	
XX	
PD 11-MAY-2000.	
XX	
PF 29-OCT-1999;	99WO-GB03579.
XX	
PR 29-OCT-1998;	98US-0106070.
PR 01-MAR-1999;	99US-0122066.
PR 27-OCT-1999;	99US-0428122.
XX	
PA (CONN-) CONNAUGHT LAB LTD.	
XX	
PI Murdin AD, Oomen RP, Dunn PL;	
XX	
DR WPI: 2000-365569/31.	
DR P-PSDB: Y94327.	

XX Novel Chlamydia 98 kDa putative outer membrane protein antigen, used  
PT for vaccination and protection against Chlamydia infection  
XX  
XX Claim 1; Fig 1; 93pp; English.

CC The present sequence is the 98kDa putative outer membrane protein gene  
CC from Chlamydia pneumoniae. The genomic sequence was amplified using two  
CC PCR primers. The 5' primer contains a NotI restriction site, a ribosome  
CC binding site, an initiation codon and a sequence close to the 5' end of  
CC the 98kDa putative outer membrane protein coding sequence. The 3' primer  
CC contains the sequence encoding the C-terminal sequence of the putative  
CC outer membrane protein and a BstXI restriction site. The stop codon was  
CC excluded and an additional nucleotide was inserted to obtain an in-frame  
CC C-terminal fusion with the Histidine tag. The PCR product was cloned  
CC into a eukaryotic expression vector (pCA-Myc-His) by restricting both  
CC the vector and the PCR product with NotI and BamHI and performing a  
CC ligation reaction. This expression vector was injected intramuscularly  
CC and intranasally into mice, which were subsequently inoculated with  
CC Chlamydia pneumoniae. The chlamydial lung titers of the immunised mice  
CC were lower than those of the controls. Thus the 98kDa putative outer  
CC membrane protein can be used as a vaccine to provide protection against  
CC Chlamydia infections, especially Chlamydia pneumoniae infections.  
CC The polypeptide may also be administered orally to treat Chlamydia  
CC infection. The present sequence may also be used in the  
CC construction of attenuated Chlamydia strains that can over-express the  
CC gene or express it in a non-toxic form.

XX Sequence 3000 BP; 871 A; 617 C; 634 G; 878 T; 0 other;

Query Match 100.0%; Score 3000; DB 21; Length 3000;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 3000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cgcctctactagtagaggttgatgaattcttgactgtttccctatgtgtatct 60  
DB 1 cgcctctactagtagaggttgatgaattcttgactgtttccctatgtgtatct 60  
QY 61 cttaaatatataatcaaaaagatatattttacaaatgaagctctctcccaa 120  
DB 61 cttaaatatataatcaaaaagatatattttacaaatgaagctctctcccaa 120  
QY 121 gtttgatatttctacattgtcatttcccttgctcattatgttcacagacagttt 180  
DB 121 gtttgatatttctacattgtcatttcccttgctcattatgttcacagacagttt 180  
QY 181 ggtttcaagtcgaggttcgattggaaataaattgttaattttcagttcgtgagatca 240  
DB 181 ggtttcaagtcgaggttcgattggaaataaattgttaattttcagttcgtgagatca 240  
QY 241 ggaagatgtcgaactacctaattttaaggaaatgtcactctgaataatttccctg 300  
DB 241 ggaagatgtcgaactacctaattttaaggaaatgtcactctgaataatttccctg 300  
QY 301 aacagcacaagcaatcaaaaaagctgttttaacaaactaaggagattgaattccac 360  
DB 301 aacagcacaagcaatcaaaaaagctgttttaacaaactaaggagattgaattccac 360  
QY 361 aggttaacgggaactctctattgttccaacggtgagtcagggacgtgacagggctgc 420  
DB 361 aggttaacgggaactctctattgttccaacggtgagtcagggacgtgacagggctgc 420  
QY 421 tgttaacacagcgttgatataatctacacagtttataggttttcttcgtactctt 480  
DB 421 tgttaacacagcgttgatataatctacacagtttataggttttcttcgtactctt 480  
QY 481 tattgctctctctggaagtcgataactacccggaaggagccgttagtctgtctacgg 540  
DB 481 tattgctctctctggaagtcgataactacccggaaggagccgttagtctgtctacgg 540  
QY 541 tagcttgagttgacaaaaatgtcagttgtctcttcacaaaacttttcaacgataa 600  
DB 541 tagcttgagttgacaaaaatgtcagttgtctctcttcacaaaacttttcaacgataa 600

DB 541 tagcttgagttgacaaaaatgtcagttgtctctcttcacaaaacttttcaacgataa 600  
QY 601 tggggggtcattacacgcgcaaaaactctttcattacaaggatacaatgtcagctgtt 660  
DB 601 tggggggtcattacacgcgcaaaaactctttcattacaaggatacaatgtcagctgtt 660  
QY 661 ttctgaataatactctcccaagaagaagcgagagccattcagactctccgctccat 720  
DB 661 ttctgaataatactctcccaagaagaagcgagagccattcagactctccgctccat 720  
QY 721 tactggaacccaagggaagtcctttttctgacaatacttcttcgagattctgagctgc 780  
DB 721 tactggaacccaagggaagtcctttttctgacaatacttcttcgagattctgagctgc 780  
QY 781 aattttacagaagcctcgtgtgctatttcttaataatgttaagttccttattgaca 840  
DB 781 aattttacagaagcctcgtgtgctatttcttaataatgttaagttccttattgaca 840  
QY 841 taagttacagagcgagctcctccacaacacgggagatatgctcaggggtctctgc 900  
DB 841 taagttacagagcgagctcctccacaacacgggagatatgctcaggggtctctgc 900  
QY 901 tttaaaactagtagacataactaagttcacccctcactggaatcagatgttactctcag 960  
DB 901 tttaaaactagtagacataactaagttcacccctcactggaatcagatgttactctcag 960  
QY 961 caacaatacatcgcacaacacgcggagagctatctatgtgtaaaaaagctcagctgttc 1020  
DB 961 caacaatacatcgcacaacacgcggagagctatctatgtgtaaaaaagctcagctgttc 1020  
QY 1021 cggagagcttaccctcttcaagtagaataatgtcgaatgaggttaacgctcctaaggtg 1080  
DB 1021 cggagagcttaccctcttcaagtagaataatgtcgaatgaggttaacgctcctaaggtg 1080  
QY 1081 agccatagctatcgaagatagtgagggaattgttatccgcccagatgtgtgacattgt 1140  
DB 1081 agccatagctatcgaagatagtgagggaattgttatccgcccagatgtgtgacattgt 1140  
QY 1141 ctttttagggaatacagtcactctcactactcctcctgagcgaataagaatagatcagact 1200  
DB 1141 ctttttagggaatacagtcactctcactactcctcctgagcgaataagaatagatcagact 1200  
QY 1201 aggaacgagtcgaagaatgacagcttgcgttcgctcgtcgtgtagagccattactcta 1260  
DB 1201 aggaacgagtcgaagaatgacagcttgcgttcgctcgtcgtgtagagccattactcta 1260  
QY 1261 tgatccataactacaggaatcatccacaacagtttacaagattcttaaaagttaagaac 1320  
DB 1261 tgatccataactacaggaatcatccacaacagtttacaagattcttaaaagttaagaac 1320  
QY 1321 tccggcagattctgcaactacaataataggaacatcactcttccacagagaagaattac 1380  
DB 1321 tccggcagattctgcaactacaataataggaacatcactcttccacagagaagaattac 1380  
QY 1381 agagcagaagcgcggaagtttcaaaatcttactctcgaagctactacagcctgttaactct 1440  
DB 1381 agagcagaagcgcggaagtttcaaaatcttactctcgaagctactacagcctgttaactct 1440  
QY 1441 ttcaagaggtactcctcatttcaaaacatgtagtgcctctgagactcagagattcacca 1500  
DB 1441 ttcaagaggtactcctcatttcaaaacatgtagtgcctctgagactcagagattcacca 1500  
QY 1501 acaggcagattctcgtctcgaatgagacgtgagactactctagaaactcgtgtactag 1560  
DB 1501 acaggcagattctcgtctcgaatgagacgtgagactactctagaaactcgtgtactag 1560  
QY 1561 caccataaacaattgtgcatataatcagttcattagacgtgacaaaggacaaat 1620  
DB 1561 caccataaacaattgtgcatataatcagttcattagacgtgacaaaggacaaat 1620  
QY 1621 agaaaccaaagctacgtcaaaaaatctgacttatacggaaacataacttattgagccc 1680  
DB 1621 agaaaccaaagctacgtcaaaaaatctgacttatacggaaacataacttattgagccc 1680

```

QY 1661 gacggcagctttatgtaaaatcagtttaagaatcctcagctcagcagcttaga 1740
    |||
Db 1661 gacggcagctttatgtaaaatcagtttaagaatcctcagctcagcagcttaga 1740
QY 1741 gctcaaaagcttcgtaactgtaacagacagcgagtgatccagatctctataaagggtga 1800
    |||
Db 1741 gctcaaaagcttcgtaactgtaacagacagcgagtgatccagatctctataaagggtga 1800
QY 1801 gaattccatcagcgctatcagcgagctggggcccaattgttgagggaaggagcttc 1860
    |||
Db 1801 gaattccatcagcgctatcagcgagctggggcccaattgttgagggaaggagcttc 1860
QY 1861 taagactgcaacccctcaactgtaactgtaaaactggtatatactcctaattcccgagcgatcgg 1920
    |||
Db 1861 taagactgcaacccctcaactgtaactgtaaaactggtatatactcctaattcccgagcgatcgg 1920
QY 1921 ctctttagctccatataagctatagtaagatgcatatataattagctctccatctatct 1980
    |||
Db 1921 ctctttagctccatataagctatagtaagatgcatatataattagctctccatctatct 1980
QY 1981 tatggaagctgcaaaacggaagggttcgcaaggagacggctgtttgtgtgctcgagatctc 2040
    |||
Db 1981 tatggaagctgcaaaacggaagggttcgcaaggagacggctgtttgtgtgctcgagatctc 2040
QY 2041 taactcttcataagaagtagtacaaaaacacgacgaggttcgcacatttgagtgagcg 2100
    |||
Db 2041 taactcttcataagaagtagtacaaaaacacgacgaggttcgcacatttgagtgagcg 2100
QY 2101 ttatgtctagaggaagaaacccacactgtgttcagatagaattctctatgtgtctatcttg 2160
    |||
Db 2101 ttatgtctagaggaagaaacccacactgtgttcagatagaattctctatgtgtctatcttg 2160
QY 2161 tcaagctctttggaagagaatagaactactgttgaactgaagaatcaaggtaacgctacag 2220
    |||
Db 2161 tcaagctctttggaagagaatagaactactgttgaactgaagaatcaaggtaacgctacag 2220
QY 2221 aggaactctctatataccgacgacaaagaaactatctctctctctctctctctctctctc 2280
    |||
Db 2221 aggaactctctatataccgacgacaaagaaactatctctctctctctctctctctctc 2280
QY 2281 ttgttcgtgtctatgtgtctctacagagatctctctctctctctctctctctctctctct 2340
    |||
Db 2281 ttgttcgtgtctatgtgtctctacagagatctctctctctctctctctctctctctctct 2340
QY 2341 caccatacggataacgaatctgtaaaacccaagtatacaacatactactgtttaagagag 2400
    |||
Db 2341 caccatacggataacgaatctgtaaaacccaagtatacaacatactactgtttaagagag 2400
QY 2401 ctggggggaatagatcttcgctttagaattcgggtggaagagctccgattctctatagta 2460
    |||
Db 2401 ctggggggaatagatcttcgctttagaattcgggtggaagagctccgattctctatagta 2460
QY 2461 aaggtcctatttgagcagatcatagcctctatagaattggaattgtgctctgacacaa 2520
    |||
Db 2461 aaggtcctatttgagcagatcatagcctctatagaattggaattgtgctctgacacaa 2520
QY 2521 ggaaggttttaagaagacgggaacagaaagctcgtaatttggaagtacgctctgtgtaa 2580
    |||
Db 2521 ggaaggttttaagaagacgggaacagaaagctcgtaatttggaagtacgctctgtgtaa 2580
QY 2581 tctgtccttactactcggagatccgatttgataaagaatcagactgcgaagtgtgcaacta 2640
    |||
Db 2581 tctgtccttactactcggagatccgatttgataaagaatcagactgcgaagtgtgcaacta 2640
QY 2641 caatcaactctgtgtatactgtgactgtgtcgtatgaaccggcgctgtacagaaac 2700
    |||
Db 2641 caatcaactctgtgtatactgtgactgtgtcgtatgaaccggcgctgtacagaaac 2700
QY 2701 actcggaattagcggtgtctcttggaaaacctcggtatgagaatttggaagaagaagcttt 2760
    |||
Db 2701 actcggaattagcggtgtctcttggaaaacctcggtatgagaatttggaagaagaagcttt 2760

```

```

QY 2761 agtctctgcaagggaaccattttgtcttaactcaaatlttgaaagccttagcaatt 2820
    |||
Db 2761 agtctctgcaagggaaccattttgtcttaactcaaatlttgaaagccttagcaatt 2820
QY 2821 ttctttgaattgcgtggtggtatctgcgaattacatgtagaacttaggaagaaatacca 2880
    |||
Db 2821 ttctttgaattgcgtggtggtatctgcgaattacatgtagaacttaggaagaaatacca 2880
QY 2881 atttaattgcttagctttgttgaaggctccacatcgaaggaaggaaagcctttaag 2940
    |||
Db 2881 atttaattgcttagctttgttgaaggctccacatcgaaggaaggaaagcctttaag 2940
QY 2941 attcttgaagcctctttcatttcgattccattcgaatttagtgtttgtctaaacacttc 3000
    |||
Db 2941 attcttgaagcctctttcatttcgattccattcgaatttagtgtttgtctaaacacttc 3000

RESULT 2
X91990
ID X91990 standard; DNA; 1230025 BP.
XX
AC X91990;
XX
DT 13-SEP-1999 (first entry)
XX
DE Nucleotide sequence of the complete genome of Chlamydia pneumoniae.
KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW vaccine; neutralising epitope; ss.
XX
OS Chlamydia pneumoniae.
XX
PN WO9927105-A2.
XX
PD 03-JUN-1999.
XX
PF 20-NOV-1998; 98WO-IB01890.
XX
PR 04-NOV-1998; 98US-0107078.
PR 21-NOV-1997; 97FR-0014673.
XX
PA (GEST ) GENSET.
XX
PI Grifais R;
XX
DR WPI; 1999-357842/30.
XX
PT Genome sequence of Chlamydia pneumoniae
PS Claim 1; Page 291-611; 1912pp; English.
XX
CC The present sequence represents the complete genome of Chlamydia
CC pneumoniae, and encodes proteins Y34584-Y35879. C. pneumoniae causes
CC respiratory disease such as pneumonia and bronchitis and is thought
CC to be a contributing factor in heart disease, sarcoidosis, sinusitis,
CC purulent otitis media, erythema nodosum or pharyngitis. The polypeptides
CC encoded by the open reading frames of the C. pneumoniae genome (see
CC Y34584-Y35879) can be used in immunogenic compositions as vaccines.
CC Vectors containing C. pneumoniae nucleotide sequences can also be
CC used as immunogenic compositions, especially where the vector directs
CC the expression of a neutralising epitope of C. pneumoniae.
XX
SQ Sequence 1230025 BP; 367213 A; 249833 C; 249013 G; 363589 T; 377 other;

Query Match 99.28; Score 2975.4; DB 20; Length 1230025;
Best Local Similarity 99.98; Pred. No. 0;
Matches 2998; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1 cgccttaccagtagaggttgagtgatcttctgactgtttctctatgtgtatct 60
    |||
Db 31658 cgccttaccagtagaggttgagtgatcttctgactgtttctctatgtgtatct 31717

```

QY	61	cttaaatatattaatccaatccaagtatatattttaacaatgaaagttcttcccca	120
Db	31718	cttaaaaatattaatccaatccaagaatataattattttaacaatgaaagttcttcccca	31777
QY	121	gttgtaattcttcacattgtctatttccctttctcagtttgctacccggaaagttt	180
Db	31778	gtttgtaattcttcacattgtctatttccctttctcagtttgctacccggaaagttt	31833
QY	181	ggattccaagtgcaggttcgatatgggaaataaaatggttaattttcagtttcgtagatca	240
Db	31838	ggattccaagtgcaggtttcgtatgggaaataaaatggttaattttcagtttcgtagatca	3189
QY	241	ggaaagcgtgcggaaactccctacatttaaggggaaatgtaactctagaanaatctctg	300
Db	31898	ggaaagcgtgcggaaactccctacatttaaggggaaatgtaactctagaanaatctctg	3195
QY	301	aacagcgccagccaatccaacaaaacgtgtttaacaacacccaaggcggaattgacttcc	360
Db	31958	aacgcgccagccaatccaacaaaacgtgtttaacaacacccaaggcggaattgacttcc	3201
QY	361	aggtaaacgggaactctctatgttccaacggtgagatgcaggagctgtaacagggtctc	420
Db	32018	aggtaaacgggaactctctatgttccaacggtgagatgcaggagctgtaacagggtctc	3207
QY	421	tgttaaacagcagcgtgtgtatagataaatctaccacaggttatagggttttcttcgatactt	480
Db	32078	tgttaaacagcagcgtgtgtatagataaatctaccacaggttatagggttttcttcgatactt	3213
QY	481	tattgcgtctcctcgaaagttcgaataactaccgcgcgaaggacccgttagctgctctacgg	540
Db	32138	tattgcgtctcctcgaaagttcgaataactaccgcgcgaaggacccgttagctgctctacgg	3219
QY	541	taacttcgaagtttaacaacaaaatgtcagtttgccttcacgacaaaactttccaacggataa	600
Db	32198	taacttcgaagtttaacaacaaaatgtcagtttgccttcacgacaaaactttccaacggataa	3225
QY	601	tggcggtgtctataccaocgcgaanaaccttcttaataacaggaactacaaatgtcaagcttgt	660
Db	32257	tggcggtgtctataccaocgcgaanaaccttcttaataacaggaactacaaatgtcaagcttgt	3231
QY	661	ttctgnaaataactcctccaagaagaagcggagacatccaagactccgatgcccatacoat	720
Db	32317	ttctgnaaataactcctccaagaagaagcggagacatccaagactccgatgcccatacoat	3237
QY	721	tactcggaaaaccaaagggggaagctctctttcttcgacaatactcttcggattccggagctgc	780
Db	32377	tactcggaaaaccaaagggggaagctctctttcttcgacaatactcttcggattccggagctgc	3243
QY	781	aatttttaagaagcctcgtgtgactattctctaataatgctaaagttcccttatgtgaca	840
Db	32437	aatttttaagaagcctcgtgtgactattctctaataatgctaaagttcccttatgtgaca	3249
QY	841	taaggttcaaggaagcgtgtcctccaacacgggggatatgtcgaaggagtggtcattctgyc	900
Db	32497	taaggttcaaggaagcgtgtcctccaacacgggggatatgtcgaaggagtggtcattctgyc	3255
QY	901	tataataactagatacagatactaaggtcaocctccaactcgynaataicagatgttactctcag	960
Db	32557	tataataactagatacagatactaaggtcaocctccaactcgynaataicagatgttactctcag	3261
QY	961	caacatacatctgcacaacacgcggggggggtctctcatgtgaaaaagcccgaaactggttc	1020
Db	32617	caacatacatctgcacaacacgcggggggggtctctcatgtgaaaaagcccgaaactggttc	3267
QY	1021	cggaggaacttaccctattccagtagaataatgtgtcaatgtgaggttaacgtccctaaaggtg	1080
Db	32677	cggaggaacttaccctattccagtagaataatgtgtcaatgtgaggttaacgtccctaaaggtg	3273
QY	1081	agccatactctcgaagtagtgggggaattgaagttacccgcgcgaatagttgtagacattgt	1140
Db	32737	agccatactctcgaagtagtgggggaattgaagttacccgcgcgaatagttgtagacattgt	3279

QY	1141	cttttgggaataacgctcaactctctactactcctctgagcagatagaagtagtatcgactt	1200
Db	32797	cttttgggaataacagctcaactctctactactcctctgagcagatagaagtagtatcgactt	32856
QY	1201	aggaacagatgacaaagatgacagcttctgtctctgtctgtgtgtatagagcatctactcta	1280
Db	32857	aggaacagatgacaaagatgacagcttctgtctctgtctgtgtgtatagagcatctactcta	32916
QY	1261	tgatcccatatctacagcgatcatccacaacagttcacagatgcttaaaagttaatgagac	1320
Db	32917	tgatcccatatctacagcgatcatccacaacagttcacagatgcttaaaagttaatgagac	32976
QY	1321	tcgcgcagatctcgccactacaatatcacaggagacatcactctccacagagaagaattatc	1380
Db	32977	tcgcgcagatctcgccactacaatatcacaggagacatcactctccacagagaagaattatc	33036
QY	1381	agagacagagcgccgcagattcttaaaatcttaacttcgagatctacacagcctgtgaactc	1440
Db	33037	agagacagagcgccgcagattcttaaaatcttaacttcgagatctacacagcctgtgaactc	33096
QY	1441	tttaaggagagactctctcttttaaacatgagagtgactcttgagactcaaggcatctacta	1500
Db	33097	tttaaggagagactctctcttttaaacatgagagtgactcttgagactcaaggcatctacta	33156
QY	1501	acagcgagatctctcgtctcgtgaatgtgagctgaggaactcctctagaacctgtatattag	1560
Db	33157	acagcgagatctctcgtctcgtgaatgtgagctgaggaactcctctagaacctgtatattag	33216
QY	1561	cacacataaaccaattggtgcatattaacatgaattctctatagacggtgcacaaagacaaaat	1620
Db	33217	cacacataaaccaattggtgcatattaacatgaattctctatagacggtgcacaaagacaaaat	33276
QY	1621	agaaacacaaagatcagtcacaaaaaactgaacttctcgagacacatcaactttatggacc	1680
Db	33277	agaaacacaaagatcagtcacaaaaaactgaacttctcgagacacatcaactttatggacc	33336
QY	1681	gaagggcacaagtttatgtgaaatcatatgattttaagaactcctcagtcctacagacatctaga	1740
Db	33337	gaagggcacaagtttatgtgaaatcatatgattttaagaactcctcagtcctacagacatctaga	33396
QY	1741	gtctcaaaagctctctggaactctgtaacaaagacccgcgtgacccagatccctataatgtgtga	1800
Db	33397	gtctcaaaagctctctggaactctgtaacaaagacccgcgtgacccagatccctataatgtgtga	33456
QY	1801	gaaatccatcacggtctacaggggaacttggtggcccaattggttgggggacaggggttc	1860
Db	33457	gaaatccatcacggtctacaggggaacttggtggcccaattggttgggggacaggggttc	33516
QY	1861	tacgcgctcacaactccaactggaactaaactggtctatatcttccaatcccgagcgtatcgg	1920
Db	33517	tacgcgctcacaactccaactggaactaaactggtctatatcttccaatcccgagcgtatcgg	33576
QY	1921	cctctttagtccctaatagtactatggaatgcatattatagatatagctctctccatatct	1980
Db	33577	cctctttagtccctaatagtactatggaatgcatattatagatatagctctctccatatct	33636
QY	1981	tatggaagctgcgcaaacgaaaggtgtgcagggagacgcgtcttthtggtygtcgtgattac	2040
Db	33637	tatggaagctgcgcaaacgaaaggtgtgcagggagacgcgtcttthtggtygtcgtgattac	33696
QY	2041	taactcttcacataagatgataaanaaacaagacgcgggttccgcgcatctgagtgagcgcg	2100
Db	33697	taactcttcacataagatgataaanaaacaagacgcgggttccgcgcatctgagtgagcgcg	33756
QY	2101	ttatgtcataggggaaacctacatactgttccagaataagattcttagtgctgcacttttg	2160
Db	33757	ttatgtcataggggaaacctacatactgttccagaataagattcttagtgctgcacttttg	33816
QY	2161	tcaactctttggaagagatagagacacttttgagtaagaatctttagtgctgcacttttg	2219
Db	33817	tcaactctttggaagagatagagacacttttgagtaagaatctttagtgctgcacttttg	33876
QY	2220	gaggaaccccttatcacaacgaaacacatatactctctcttccctgcgaacatctacgc	2279

```

Db 33877 gaggaactctctattaccagcacaacgaacatactctctctcctcgaacacagc 33936
Qy 2280 cttgtctgttgcctatcttctctagaagattcctgttctcttccaggaacctta 2339
Db 33937 cttgtctgttgcctatcttctctagaagattcctgttctcttccaggaacctta 33936
Qy 2340 aacaccatacggataacacatcgaacaaacgaatatacacaatctctcgtttaaggaa 2399
Db 33997 aacaccatacggataacacatcgaacaaacgaatatacacaatctctcgtttaaggaa 34036
Qy 2400 gctggggaatgatagttcgttctcgtttaagaattcgttgaagagctcgtttgtagatg 2459
Db 34057 gctggggaatgatagttcgttctcgtttaagaattcgttgaagagctcgtttgtagatg 34116
Qy 2460 aagagtcctattgagagatcagtcacccctcctcctcctcctcctcctcctcctc 2519
Db 34117 aagagtcctattgagagatcagtcacccctcctcctcctcctcctcctcctcctc 34176
Qy 2520 aggaaggttttaagaacagaggaacagaagctcgttgaatttgaagtagccgtctgtga 2579
Db 34177 aggaaggttttaagaacagaggaacagaagctcgttgaatttgaagtagccgtctgtga 34236
Qy 2580 atctgcttctcctcctcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 2639
Db 34237 atctgcttctcctcctcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 34296
Qy 2640 acaactactcttctgttatactgtgactcgttgcgttgcgttgcgttgcgttgcgttgcgt 2699
Db 34297 acaactactcttctgttatactgtgactcgttgcgttgcgttgcgttgcgttgcgttgcgt 34356
Qy 2700 cactgcgaatagcgggtgactccttgcgaacacctcgttgaagatttgcgaagacagctt 2759
Db 34357 cactgcgaatagcgggtgactccttgcgaacacctcgttgaagatttgcgaagacagctt 34416
Qy 2760 tagtctcctcgttgcgaaggaacaccttcttcttaactaaatttgaacaccttgaacacct 2819
Db 34417 tagtctcctcgttgcgaaggaacaccttcttcttaactaaatttgaacaccttgaacacct 34476
Qy 2820 ttctcttgaattgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 2879
Db 34477 ttctcttgaattgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 34536
Qy 2880 aattctaatggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 2939
Db 34537 aattctaatggttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 34596
Qy 2940 gattcctgaagagctccttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 2999
Db 34597 gattcctgaagagctccttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 34656
Qy 3000 c 3000
Db 34657 c 34657

```

```

PD 30-DEC-1998.
XX
PF 19-JUN-1998; 98NO-DK00266.
XX
PR 23-JUN-1997; 97DK-0000744.
XX
PA (BIRK/) BIRKELUND S.
XX (CHR1/) CHRISTIANSEN G.
PI Birkelund S, Christiansen G, Knudsen K, Madsen A;
PI Mygind P;
DR WPI: 1999-105610/09.
DR P-PSDB; W88421.
XX
PT Species-specific test for identifying mammals infected with
PT Chlamydia pneumoniae - comprises detecting antibodies specific for
PT outer membrane proteins of C. pneumoniae or nucleic acids encoding
PT these proteins
PS Claim 6; Page 52-53; 115pp; English.
XX
CC This DNA sequence codes for the novel 90.0 kDa surface exposed
CC protein Omp8 (see W88421) of the human respiratory pathogen
CC Chlamydia pneumoniae. By generating antibodies against C.
CC pneumoniae outer membrane complex, a polyclonal antibody (PAB 150)
CC was obtained which reacted with outer membrane proteins. The
CC antibody was used to identify the genes (see X06816-27) encoding
CC Omp4-Omp5 proteins (see W88417-28) in an expression library of
CC C. pneumoniae DNA. The genes are situated in 2 gene clusters:
CC Omp12,11,10,5,4,13 and 14 in one cluster and Omp6,7,8,9 and 15 in
CC the other, and encode polypeptides of about 89.6-100.3 kDa and
CC about 56.1 kDa. The invention provides a new species specific test
CC for identifying mammals (including humans) infected with Chlamydia
CC pneumoniae. The test comprises detecting antibodies specific for
CC Omp4-Omp5 or detecting nucleic acid fragments encoding these outer
CC membrane proteins, especially by PCR. The proteins are also used
CC in the diagnosis of C. pneumoniae infection in mammals. The
CC nucleic acids and proteins can also be used in the immunization of
CC mammals, the nucleic acids being particularly useful as DNA
CC vaccines for effecting in vivo expression of antigens. The
CC vaccines may also prevent atherosclerosis and bronchial asthma,
CC which are possibly associated with C. pneumoniae.
XX
SQ Sequence 2787 BP; 811 A; 583 C; 598 G; 795 T; 0 other;

```

Query Match 92.5%; Score 2775.8; DB 20; Length 2787;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 2780; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

Qy 101 atgaagctctctccccaagttgtattctcacttgcatttcccttgcattg 160
Db 1 atgaagctctctccccaagttgtattctcacttgcatttcccttgcattg 60
Qy 161 atgctacccgagacagtttggatccaagtgcgagtttcgtagggaaataatgta 220
Db 61 atgctacccgagacagtttggatccaagtgcgagtttcgtagggaaataatgta 120
Qy 221 ttctcaattggtgagtgagtgaggaagatgcgtggaactcctcactatttaaggaaatg 280
Db 121 ttctcaattggtgagtgagtgaggaagatgcgtggaactcctcactatttaaggaaatg 180
Qy 281 actctagaaaatattcctcgtgaacagcgacagcaatcacaagaagcttttaacaact 340
Db 181 actctagaaaatattcctcgtgaacagcgacagcaatcacaagaagcttttaacaact 240
Qy 341 aagggcatttgacttccacaggaacggaactctatgttccaaagcgtgagatga 400
Db 241 aagggcatttgacttccacaggaacggaactctatgttccaaagcgtgagatga 300
Qy 401 gggactgtacaggggctgtgtttaacagcagcgtgtgataataattaccacgttata 460

```



```
Db 301 gggactgtacagggctgctgtttaacagcgtgtgtataaatctacacggttata 360
QY 461 gggtttcttcgctatcctttatctgtctctcgtgaagttcgtatacctacggcaagg 520
Db 361 gggtttcttcgctatcctttatctgtctctcgtgaagttcgtatacctacggcaagg 420
QY 521 ggcgttaagcgtcctcgggttagctgtgattgacaaaaaagtcagttgtcttcagc 580
Db 421 ggcgttaagcgtcctcgggttagctgtgattgacaaaaaagtcagttgtcttcagc 480
QY 581 aaaaactttccacaggaataatggcgtgtatcacccgcaaaacctttcatatacagg 640
Db 481 aaaaactttccacaggaataatggcgtgtatcacccgcaaaacctttcatatacagg 540
QY 641 actacatctcagctctgtttctcgaataaacctccctcaaaaaggcgagccattcag 700
Db 541 actacatctcagctctgtttctcgaataaacctccctcaaaaaggcgagccattcag 600
QY 701 acttcgcagtccttacccttactctgtaaaccaagggaagttctttctcgcacatact 760
Db 601 acttcgcagtccttacccttactctgtaaaccaagggaagttctttctcgcacatact 660
QY 761 tcttcgagatctcggagcgtcgaatttttaacagaagcctcgtgtactttcctataatgct 820
Db 661 tcttcgagatctcggagcgtcgaatttttaacagaagcctcgtgtactttcctataatgct 720
QY 821 aagatttccttattgacaaataaggltcacagagcgagctcctcaacacgggggataatg 880
Db 721 aagatttccttattgacaaataaggltcacagagcgagctcctcaacacgggggataatg 780
QY 881 tcaggaaggtgcctatctgtgtctataaaaactagtaacagataggttcaaccttactgga 940
Db 781 tcaggaaggtgcctatctgtgtctataaaaactagtaacagataggttcaaccttactgga 840
QY 941 aatcgaaggttacctcttcaagaacaatacatcgcgaacacgggagagctatctatgtg 1000
Db 841 aatcgaaggttacctcttcaagaacaatacatcgcgaacacgggagagctatctatgtg 900
QY 1001 aaaaagctcgaactggtctcggaggaacttaccctattcagtagaataatgttcaatgga 1060
Db 901 aaaaagctcgaactggtctcggaggaacttaccctattcagtagaataatgttcaatgga 960
QY 1061 ggtacagctccttaagggttgagcgccatagctatcgaagaatgttggaatttatactc 1120
Db 961 ggtacagctccttaagggttgagcgccatagctatcgaagaatgttggaatttatactc 1020
QY 1121 gccgatagtgtgacatgttcttttagggaaatacagttcaactcttactctctggagc 1180
Db 1021 gccgatagtgtgacatgttcttttagggaaatacagttcaactcttactctctggagc 1080
QY 1181 aatagaagttgatatcgacttagaagaacgagtgcagaagaatgacagcttgcgtctgctc 1240
Db 1081 aatagaagttgatatcgacttagaagaacgagtgcagaagaatgacagcttgcgtctgctc 1140
QY 1241 ggtagaagccatctactctctatgtatccataactacagaggttcatcccaacagttacagat 1300
Db 1141 ggtagaagccatctactctctatgtatccataactacagaggttcatcccaacagttacagat 1200
QY 1301 gttcttaaaagttaatgagactccgagcagattctgcacatcaatatcacagggaactcac 1360
Db 1201 gttcttaaaagttaatgagactccgagcagattctgcacatcaatatcacagggaactcac 1260
QY 1361 ttacagggagaaagttaatacagagacagagggccgagattcttaaaatcttaacttcgaag 1420
Db 1261 ttacagggagaaagttaatacagagacagagggccgagattcttaaaatcttaacttcgaag 1320
QY 1421 ctactacacgctgttaactcttcaagaggttactctatctttaaacaatggaatgtgctctg 1480
Db 1321 ctactacacgctgttaactcttcaagaggttactctatctttaaacaatggaatgtgctctg 1380
QY 1481 cagactcagagcattcaactaacagagcagattctgctctgaagaatgagcgttagaactact 1540
Db 1381 cagactcagagcattcaactaacagagcagattctgctctgaagaatgagcgttagaactact 1440
QY 1541 ctgaacacctgtactactagcacacataaacaatttgttcaatcaatcagttctatagac 1600
Db 1441 ctgaacacctgtactactagcacacataaacaatttgttcaatcaatcagttctatagac 1500
QY 1601 ggtcacaagaagggcaaaaataagaacaacgaagctacgttcaaaaaatctgaacttatctgga 1660
Db 1501 ggtcacaagaagggcaaaaataagaacaacgaagctacgttcaaaaaatctgaacttatctgga 1560
QY 1661 accatcatttatgtgaccccgagcggtcagcagtttataagaaatcatatagtttaagaaatcct 1720
Db 1561 accatcatttatgtgaccccgagcggtcagcagtttataagaaatcatatagtttaagaaatcct 1620
QY 1721 cagtcctcagacattcttaagctcctcaagccttctggaactgttaacaagcccgagttgact 1780
Db 1621 cagtcctcagacattcttaagctcctcaagccttctggaactgttaacaagcccgagttgact 1680
QY 1781 ccagatcctataatgggttggaatccatctacggtatcagggaacttggggcccaatt 1840
Db 1681 ccagatcctataatgggttggaatccatctacggtatcagggaacttggggcccaatt 1740
QY 1841 gtttgggggacaggggcttctacgactgcaaccttcaactggaactaaactgtctatat 1900
Db 1741 gtttgggggacaggggcttctacgactgcaaccttcaactggaactaaactgtctatat 1800
QY 1901 cctaatcccgagaggtatcggtcctttagtccttaataagcttatggaatgcatattagat 1960
Db 1801 cctaatcccgagaggtatcggtcctttagtccttaataagcttatggaatgcatattagat 1860
QY 1961 attagctctccatctatcttattggaactggaacggaagggttcagggagacggctct 2020
Db 1861 attagctctccatctatcttattggaactggaacggaagggttcagggagacggctct 1920
QY 2021 ttctgtgtgctgggttttcttaactcttccataaggaatgtacaaaaacagacggcg 2080
Db 1921 ttctgtgtgctgggttttcttaactcttccataaggaatgtacaaaaacagacggcg 1980
QY 2081 ttccgcaatttgaggggttatgttcatagagaagaacccatacttgttctagataag 2140
Db 1981 ttccgcaatttgaggggttatgttcatagagaagaacccatacttgttctagataag 2040
QY 2141 attcttagtgcctgcaatttctgcagctctctggaagagaatagactattgttagtaag 2200
Db 2041 attcttagtgcctgcaatttctgcagctctctggaagagaatagactattgttagtaag 2100
QY 2201 aatcaagggtacagttacagtgaggaactctctattacacagacagaaacctatctct 2260
Db 2101 aatcaagggtacagttacagtgaggaactctctattacacagacagaaacctatctct 2160
QY 2261 ctctcttgcacaactcagggccttgcgtgtgtctatgttctctacagagattcctgtctc 2320
Db 2161 ctctcttgcacaactcagggccttgcgtgtgtctatgttctctacagagattcctgtctc 2220
QY 2321 ttcttaagaaacctttagctatacccatagagataagatttgaaaacagatatacaaca 2380
Db 2221 ttcttaagaaacctttagctatacccatagagataagatttgaaaacagatatacaaca 2280
QY 2381 tatcctactgtttaaaggaaagctgggggaatgatatgttgcgtttagaatcgggtggaaga 2440
Db 2281 tatcctactgtttaaaggaaagctgggggaatgatatgttgcgtttagaatcgggtggaaga 2340
QY 2441 gttccgatttggcttaagtgaagtgtctctatttgagcaggtatcatgccccttcagaattg 2500
Db 2341 gttccgatttggcttaagtgaagtgtctctatttgagcaggtatcatgccccttcagaattg 2400
QY 2501 cagtttctatgcacatcagaggaaggtttaaagaaacagggaaacagaaactcgtgtaatt 2560
Db 2401 cagtttctatgcacatcagaggaaggtttaaagaaacagggaaacagaaactcgtgtaatt 2460
QY 2561 ggaagtagccggtctgtgaaatcttgccttaactatcgggacgcagattgtataaggaata 2620
Db 2461 ggaagtagccggtctgtgaaatcttgccttaactatcgggacgcagattgtataaggaata 2520
```

```

QY 2621 gactgcacagatgacgtacacatcaactctgtgtatctgtgtgagatctgttcgtaagt 2680
DB 2521 gactgcacagatgacgtacacatcaactctgtgtatctgtgtgagatctgttcgtaagt 2580
QY 2661 aacccgactctacgacacacactgcgaattgagcggtgattcttggaaaaccttcgtaag 2740
DB 2581 aacccgactctacgacacacactgcgaattgagcggtgattcttggaaaaccttcgtaag 2640
QY 2741 aattggcaaacgaacgcttgccttcctgcgcggaagcaaccttcttctaactcaat 2800
DB 2641 aattggcaaacgaacgcttgccttcctgcgcggaagcaaccttcttctaactcaat 2700
QY 2801 ttgaaagcctttagcgaattcttcttgaattgctgttgatctcgcgaattacaatgta 2860
DB 2701 ttgaaagcctttagcgaattcttcttgaattgctgttgatctcgcgaattacaatgta 2760
QY 2861 gacttagagagaataatcaattctaa 2887
DB 2761 gacttagagagaataatcaattctaa 2787

RESULT 4
X06817 standard: DNA; 2815 BP.
XX
XX X06817;
XX
XX 26-APR-1999 (first entry)
XX
XX Chlamydia pneumoniae surface exposed protein Omp5 DNA.
XX
XX Omp5: outer membrane protein 5; surface exposed protein; antigen;
XX infection; diagnosis; vaccine; atherosclerosis; asthma; ss:
XX Chlamydia pneumoniae.
XX
XX Key Location/Qualifiers
XX CDS 1..2787
XX FT /*tag= a
XX
XX MO9858953-A2.
XX
XX 30-DEC-1998.
XX
XX 19-JUN-1998; 98WO-DK00266.
XX
XX 23-JUN-1997; 97DK-0000744.
XX
XX (BIRK/) BIRKELUND S.
XX (CHRI/) CHRISTIANSEN G.
XX
XX Birkelund S, Christiansen G, Knudsen K, Madsen A;
XX Mylind P;
XX WPI: 1999-105610/09.
XX P-PSDB: W88418.
XX
XX Species-specific test for identifying mammals infected with
XX Chlamydia pneumoniae - comprises detecting antibodies specific for
XX outer membrane proteins of C. pneumoniae or nucleic acids encoding
XX these proteins
XX
XX Claim 6; Page 42-43; 115pp; English.
XX
XX This DNA sequence codes for the novel 97.2 kDa surface exposed
XX protein Omp5 (see W88418) of the human respiratory pathogen
XX Chlamydia pneumoniae. By generating antibodies against C.
XX pneumoniae outer membrane complex, a polyclonal antibody (PAB 150)
XX was obtained which reacted with outer membrane proteins. The
XX antibody was used to identify the genes (see X06816;27) encoding
XX Omp4-Omp5 proteins (see W88417-28) in an expression library of
XX C. pneumoniae DNA. The genes are situated in 2 gene clusters:
XX Omp12,11,10,5,4,13 and 14 in one cluster and Omp6,7,8,9 and 15 in

```

```

CC the other, and encode polypeptides of about 89.6-100.3 kDa and
CC about 56.1 kDa. The invention provides a new species specific test
CC for identifying mammals (including humans) infected with Chlamydia
CC pneumoniae. The test comprises detecting antibodies specific for
CC Omp4-Omp5 or detecting nucleic acid fragments encoding these outer
CC membrane proteins, especially by PCR. The proteins are also used
CC in the diagnosis of C. pneumoniae infection in mammals. The
CC nucleic acids and proteins can also be used in the immunisation of
CC mammals, the nucleic acids being particularly useful as DNA
CC vaccines for effecting in vivo expression of antigens. The
CC vaccines may also prevent atherosclerosis and bronchial asthma,
CC which are possibly associated with C. pneumoniae.
XX
XX Sequence 2815 BP; 790 A; 575 C; 626 G; 824 T; 0 other;

Query Match 17.7%; Score 530.2; DB 20; Length 2815;
Best Local Similarity 52.9%; Pred. No. 7.9e-137;
Matches 1415; Conservative 0; Mismatches 1183; Indels 75; Gaps 10;

QY 247 tgcttgacactactactatttaagggaatgcatctctgaataatctcgtgaacag 306
DB 156 tactggaatgactatactctgcagagagatataactctgcaaaccttgggattcgg- 214
QY 307 cacaagcatcacaaaagcgtgttttaacaacactaaggcgcatgttgaattcacaggtaa 366
DB 215 --cagcttaacgaaaggtgttcttcgcacactaagcaactttaaagcttcggttaa 272
QY 367 cgggaactctatctgttccaacggtgagtgcaaggactgtagcaggggtcgtttaa 426
DB 273 ggggtactcacttcttctttaaataatgaactagctcgtgaagcgacgactctcgt 332
QY 427 cagcaggtgtgtgataaactacacaggttaagggtttcttcgtactctttatgc 486
DB 333 tacaactgtataaaatcgtcgtcaacaggaattctcgagcttactcttctagcgccc 392
QY 487 gtctctggaagttcgataactaccgcgcaaggagccgttagctgtcctcaggttagctt 546
DB 393 atatcgtaatacacaaacccctcaggaaggtgagtaaatgtgaggaggtatc---t 449
QY 547 gatttgacaaaaatgctgctgtcttcagcaaaaactttcaacgataatgacg 606
DB 450 taactatgataaactggaactatatttatttaacaagattactgtgagaaatgacg 509
QY 607 tgcataccgcgaaaaactcttcatcaacagggactacaagtgcgcctgtgtctga 666
DB 510 agccattctcaccgaagaactcttcttgaaaaacagcagcgatcttctttgaag 569
QY 667 aataa-----cctcctcaagaagcgagacatcagactcgcgtcccttaag 717
DB 570 gaataaatcgagcgcaacaggaagagtggtgtattgtgcacacgtgactgtaga 629
QY 718 cattactggaaccaaagggaagtccttcttcgcacatactctcgtgattcgtgacg 777
DB 630 tattacaataataacgctctcctcctctcgcacataatgtcgaagctgcaggtg 689
QY 778 tgcatttttaacgaagccctcggtgactatttctaataatgctcaaatgttccttatta 837
DB 690 agtataataatagacaggaacgtgtacaattaaaggaataagctctgtattcttga 749
QY 838 caataaggtcacagagcgagctcctcaacaacggtgatatgtcagaggtgtactcg 897
DB 750 aataatgtgac---agcgacgcgaggaatgtgagagccttctcgtgagatgcg---- 802
QY 898 tggttataaacagtaacgataactaagttaaccccatcgtgaatcaatgttactctt 957
DB 803 -----atgttaccatctgggaatcacaagtgttaactt 836
QY 958 cagaacaataacacacagcagggagaggtatctatgttgaaaaagctcgaactggc 1017
DB 837 ctacgaagaacacagcgtgtatgtatgtcgagacattatgtcgaagaagacttaacatgcg 896
QY 1018 ttccggaagagact-----tacccattcagtagaataatgttcaatgtgagtagcagctcc 1071

```

Db	897	ttccggggggggggggtaccccttcttaacaataaagttacacatgcag	956
QY	1072	taaaagtgcgaataagatcatcagaagatagtcgggaattgattatccgcgaatgtg	1131
Db	957	taatggtggagccattctatactgcgaatcggtggagtggtgtcttccagaagaacag	1016
QY	1132	tgcagctgttttttaaggataaagtcacactcttactactc---tggacgaataag	1188
Db	1017	ggacatcaacctcoaatggaatgacatgtgctatgtgcacataccacaactacaagaagaa	1076
QY	1189	taatlacacttaaggaaacgaatgccaagaatgacagcttggcttcgtctgtgttagagc	1248
Db	1077	ttctatgacataagatgactactactcgtgaagaatcacaagaattacgtgcattatctggcatag	1136
QY	1249	cactactcttaagatcccaataacataagatacacaacagattacaagtcttaaa	1308
Db	1137	cactctttctacgatccgatctactgtataatagcgtgcagatctacagatactttaa	1196
QY	1309	agttaatggaaccgcggagatctctgacacaaataaaggggaactcaatcttcaag	1368
Db	1197	ttccaataagcgtgcatgcatgaatagtaacagatlaaagtgggcgattgttttcttgcg	1256
QY	1369	agaaaagttaacgagacagaagccgacagattcttaaaacttactctgcgaatctaca	1428
Db	1257	tgaaaagctctctgaaagatgaacaaagtgtgcagacacactcaactcttaacgca	1316
QY	1429	gctgtgaacctcttcggagagatctctatctttaaacaatgaggtgactctgcagctca	1488
Db	1317	gctgtgaacctcttaactgcgaagaaatlaagtaactaaacgtgtgcaactctgcagaa	1376
QY	1489	ggcattcaactcaacagcagatctctgctcgaatgtgacgtgaactcaactctcagaac	1548
Db	1377	agacttactcagaaccggtgtccctctgtattatgatatgacggtgcacaacgttaaaagc	1436
QY	1549	---tgctgtactcagacacataaacatattgtgcattacaatgaattctatagaacgtgc	1605
Db	1437	aagtaacagggaggttaactttaacaggtcttccactctgtgaactctttaagcgaag	1496
QY	1606	aaagaagggcaaaatagaacccaagaactcgtcaaaaactctgacttacttcgaaacct	1665
Db	1497	taagaaggtgtaattgctgcgtctctgcagcaagtaaaaatgtgcacctagvtgtccga	1556
QY	1666	caactattggaccgcgacgggcagctttatgaaaaatcagtttaagaatctcagtc	1725
Db	1557	tcctctcttggtaaaccaaggggaatgcttatgaaaatccagacttagaanaaactcaga	1616
QY	1726	ctacgacacttagagctcaaaagctctctggaactctgtaaaagaacgcgactgtactcaga	1785
Db	1617	ctttcatttgcgcagctctctgcctctgttaactgcgaacacaactacagatgttccagcgt	1676
QY	1786	tcctataatggtgtgagaatctccatcagcgtitacaggggaacttggggccaatttttg	1845
Db	1677	tcctacagtagaacctccatacgaactatggtatacgaagtacttgggaatgaactgggt	1736
QY	1846	gg-----ggacgggggctctcgcagtgaacactccaactgactgactcaaaactg	1893
Db	1737	tgaatataccgaaagcaaccccaagaactaagaacaggaatattagcttgcacaatacag	1796
QY	1894	ctatatccataaccgaagcgtatcgacactttagtccctaataagcttatggaaatgatt	1953
Db	1797	ctactctcgaactcctcgagcgtcgaagaccttagtccataagacttggggacttt	1856
QY	1954	tataaattagctctctccatatacttatggaagactcaaacgaagggtgcagggaga	2013
Db	1857	ttcagaacaccaaaggaattccaaggtctcaataagagaagtgtcttgacacttgttcaga	1916
QY	2014	ccgtctctttagtctgtgcgtatctcgaactcttcccaataaggtatgacaaaaaacg	2073
Db	1917	tcagagctcttcggtgcgtggagctgcgccaattcttagataaagaaagaaggggaaaa	1976
QY	2074	acggcggttcggcaattgagttgacgttatgtctatagtaggaaaaactacatactgttc	2133

Db	1977	acgcaataacgcgtcataataatctggttgatataatgcatatgcgaagtgccagcgcaaatctgttc	2036
QY	2134	agaataagattcttaagtcgtcgaattttgcagctctcttgaaagagataagactacttgt	21939
Db	2037	tgaaaacttaattagcttctgtcctttgcacactcttttgagsgataaagattcttctgt	20966
QY	2194	agtaagaatcaaggtataagctcctacggagaactctctattaccagcaacaaactta	22535
Db	2097	cgcataaaatcatctagataccttatcggaagcgctctatatcatcaacacattacgaagt	21566
QY	2254	tatctctctccttgcaaacacagcgcttctgtgtcttaagtcttcaagagattcc	22133
Db	2157	tagtgggttcataagttgtctccttaagtaa-----ctctcgtccttggatgcataaac	22133
QY	2314	tgctctctttcaagaaaccttagctacacccatacagataacagatctgaaacaagta	23733
Db	2214	ccctcgtttttagaagggcagcctgcgtcttatagcacgcagtaatgatctgaagaagta	22733
QY	2374	tacaacatctctactggttaaggaagcttgagggaattgatatgttcgctttagaattcg	24333
Db	2274	tactcgcatactcccgagcgtcgaaggtctcttgagggaataatgctttaacatgattg	23333
QY	2434	tggaaagacgtccgaatttgctataagatggtctcatcttgagcagatcatgcccctat	24933
Db	2334	agctctctctcatctcttaactctgataccgcgaattgttttagtaactatgctccataat	23933
QY	2494	gaaatgcagttgttctctatgcacatcagaagaagttttaagaacaggaagaagctcg	25533
Db	2394	caaacctgaaatctgcacataacgctcagacagactctctcgggaagaaggtacagaagag	24533
QY	2554	tgaatttggaaatgcgcgctctgtggaactctgctctaccatcggatcggatctgataa	26133
Db	2454	accttttgatgacgaacacctcttaactattatcttggcttaaggggtgaaagttgagaa	25133
QY	2614	ggaatcgaactgcgaagaatgcgaacgtacaactaactcttggttatactgtggaactgt	26733
Db	2514	gtctctctatgttatacttacttcttatatgatactgcgacttatctatcctatgctcgtctat	25733
QY	2674	tgtgtgaaccccgacgtctgcgaacacactgcgaatttagcggtgttcttggaaacctt	27333
Db	2574	ccgcaatgatcccaaatgcatcactacgaactgttaacagcggagcctcttggaaactta	26333
QY	2734	cgttgcgaatttggcaagaagaactttagtctctcgtgcgaaggaacaaatttggcttaa	27933
Db	2634	tgcacaataacttagcaagacagcgcttggaagtgcgtgcagcagctactacgcctctc	26933
QY	2794	ctcaaatttgaagcctttagccaatttctcttcttgaattgcgttggtatctcgaatta	28533
Db	2694	tccatgttttgaagtgcgcgcgaagtttcttcttgaagtcgtgatccctcagattta	27533
QY	2854	caatgtagacttagagcaaaatcaaatctta	286
Db	2754	taatgtagactcttgggggtgaagtccaattcca	2786
RESULT	5		
ID	261509		
XX	261509	standard; DNA; 2957 BP.	
AC	261509;		
XX	19-JUN-2000	(first entry)	
DT			
XX			
DE		DNA encoding the CPN100395 polypeptide.	
XX			
KW	CPN100395; Chlamydia infection; immune response; vaccine; ss.		
XX			
OS	Chlamydia pneumoniae.		
XX			
Key		Location/Qualifiers	
FT	101...2857		
FT	/tag= a		



```

Db 1655 gtagtgcacggggaactattatcagcatcaatctcgaacacagagcttctcct 1714
QY 1733 abcttagagctcaaaagctctctggaactgttaacaagcacccgagctagctccagatctata 1792
Db 1715 ttaataagagcttcttcgaacaaggaagcagatgactactacagatatacccgatatacccaatt 1774
QY 1793 atgggtgagaattccattacgagcatacaggaacttgg-----gcccacattgtttgg 1846
Db 1775 ctaactactcgaatcactatctggtatcagaagggaacttgaatattgtttgttgtagcagat 1834
QY 1847 gggacaaggagctctcagcagctcaacctcgaactgaactgaataaaccggtatatactcctat 1906
Db 1835 gcaactgcaaaaaaacaataaagctcctccttaacttgaactgaataaacaagagatcacagcgat 1894
QY 1907 cccgagcgatcgcgtctttagctccctaataagctatgtaagatgcatatagatataagc 1966
Db 1895 ccaggaacgctcagagccttctgtctctaataagcctgtgggtcttcttctgtcgaigtccgc 1954
QY 1967 tctcccaattatctatgagagactcgaacgaagaggttcgcggggagaccgcttcttgg 2026
Db 1955 tccattcagagccctcattgagaccggagcaaggttcgtatcttcgtcaacaatttgg 2014
QY 2027 tgtgtgattatctaaactctcctccataaaggatagatacaaaaacagacgcgggttccgc 2086
Db 2015 gtacaggaatcgcgagcttcttgcataagatcagcaagaagaaacaacagctagttatcgt 2074
QY 2087 cattgagtgcggtatctgcatagaggaagaaacctatactgttccagataagattctt 2146
Db 2075 catttaagcgggtttatgcatagagagagattcttcacgctctcgtgaataatttctt 2134
QY 2147 agtgcgcatcttgcacgtcttcttggaagagatagagctactcttgttcgcaaaataca 2206
Db 2135 aatttgccttcttcgacgtcttcttgccacgaagaagaccctctgtggtcgaagaacacat 2194
QY 2207 ggtacagctcagcggaagactctctatccagcaacaacgaactatctcttctcct 2286
Db 2195 acccatgatatgaggggagcatgagttacgcgacaccccgagaggtcctaaacccctgcct 2254
QY 2267 tgcacaactacgagcctgtctgtctgtctatgtctcctcagaagattcctgtctcttca 2326
Db 2255 aagat-----ttgtcagaagaattctgaactccctacttctgtcttcaat 2299
QY 2327 ggaaccttaagctacaccacatcagatagcatctgtaaaacagatatacaactatcct 2396
Db 2300 gctcgtgttcttattgtgcataccgacataactgaacacaagaatacagctgctatctc 2359
QY 2387 actgttaaggaagctggtggggaatgatagttctcgcttaagaattcgggtggagaagctcgc 2446
Db 2360 cctgttaagggaagctggtgggaatgatagtcctcgctgataagattgtgagagctatccgc 2419
QY 2447 a-----ttgcttagatgaagaagctctatttgagcagatcacatgcccctcctcaagaatttcag 2503
Db 2420 gtattgtcttcagagcgtcggtctcgtgtggtatccacacgcgcatttcttaaaccttag 2479
QY 2504 ttgtctatgcacatcaggaaggttttaagaacaggaaggaacagagctcgtgaatttggga 2563
Db 2480 atgctatgcacatcaggaatgactttaaggaagaaacgcacagagagcgttcttccaa 2539
QY 2564 agtgcgcgtcttgcgaatcttgccttaactcctacatccgagatccgattgataagaatcagac 2623
Db 2540 agtgaagacctctcactcactcagcgttccctgtaggaataaatttggaa-----attc 2593
QY 2624 tgcagaagtgaacatgaacttaactcgtgtatatactgtgatactgttctgttagtaaac 2683
Db 2594 tccgataagctcactagctatgactctccatagctcagcttcccgatgtgacttaagat 2653
QY 2684 ccgactgttacgaacaacatgcgaattagagctgtagcttgcgaacacccctcgtagcaaat 2743
Db 2654 ccaggtcgcagcaaaccttaigtgttccggtgagatcttgcgcgaatgtgttacaagc 2713
QY 2744 ttggcaagaagaagctttagctcctgtgtcaggggaacatttgccttaactcaaatltt 2803

```

```

Db 2714 ttgctagacaagctcttcttgcagtcgtggaatcatcatgcttgccttcaacttt 2773
QY 2804 gaaccccttaagcaattcttcttgaattcgtgtggcagctatcgcaattcaatgttagac 2863
Db 2774 gaatttccactcagcttgaagtcagatgcagaggtctctcctgtagctatgctatcagat 2833
QY 2864 ttaagagaacaataccacttcaatgc 2890
Db 2834 ctggaaggaagattcggattttaatcc 2860

RESULT 6
X06821
ID X06821 standard; DNA; 2757 BP.
XX
AC X06821;
XX
DT 26-APR-1999 (first entry)
XX
DE Chlamydia pneumoniae surface exposed protein Omp9 DNA.
XX
KW Omp9; outer membrane protein 9; surface exposed protein; antigen;
KW infection; diagnosis; vaccine; atherosclerosis; asthma; ss.
XX
OS Chlamydia pneumoniae.
XX
PN W0858953-A2.
PD 30-DEC-1998.
PF 19-JUN-1998; 98WO-DK00266.
PR 23-JUN-1997; 97DK-0000744.
XX
PA (BIRK/) BIRKELUND S.
PA (CHRT/) CHRISTIANSEN G.
PI Birkelund S, Christiansen G, Knudsen K, Madsen A;
PI Mygind P;
DR WPI: 1999-105610/09.
DR P-PDB: W06822.
XX
PT Species-specific test for identifying mammals infected with
PT Chlamydia pneumoniae - comprises detecting antibodies specific for
PT outer membrane proteins of C. pneumoniae or nucleic acids encoding
PT these proteins
XX
PS Claim 6: Page 55-56; 115pp; English.
XX
CC This DNA sequence codes for the novel 96.7 kDa surface exposed
CC protein Omp9 (see W884422) of the human respiratory pathogen
CC Chlamydia pneumoniae. By generating antibodies against C.
CC pneumoniae outer membrane complex, a polyclonal antibody (PAB 150)
CC was obtained which reacted with outer membrane proteins. The
CC antibody was used to identify the genes (see X06816-27) encoding
CC Omp4-Omp15 proteins (see W88417-28) in an expression library of
CC C. pneumoniae DNA. The genes are situated in 2 gene clusters:
CC Omp12,11,10,5,4,13 and 14 in one cluster and Omp6,7,8,9 and 15 in
CC the other, and encode polypeptides of about 89.6-100.3 kDa and
CC about 56.1 kDa. The invention provides a new species specific test
CC for identifying mammals (including humans) infected with Chlamydia
CC pneumoniae. The test comprises detecting antibodies specific for
CC Omp4-Omp15 or detecting nucleic acid fragments encoding these outer
CC membrane proteins, especially by PCR. The proteins are also used
CC in the diagnosis of C. pneumoniae infection in mammals. The
CC nucleic acids and proteins can also be used in the immunization of
CC mammals, the nucleic acids being particularly useful as DNA
CC vaccines for effecting in vivo expression of antigens. The
CC vaccines may also prevent atherosclerosis and bronchial asthma,
CC which are possibly associated with C. pneumoniae.
XX
XX Sequence 2757 BP; 770 A; 564 C; 642 G; 781 T; 0 other:

```

Query Match 17.4% Score 522.2: DB 20: Length 2757:  
Best Local Similarity 53.4%: Pred NO. 1.3e-134:  
Matches 1424: Conservative 0: Mismatches 1153: Indels 90: Gaps 12:

```
QY 242 gaagatgctggaactactactactathtaagggaatgtcactctagaaataatctctga 301
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 160 gbtgctagtggaacccattatcttcgattgggagctcgata--agccaagaggg 216
QY 302 aacagacagcaatcccaaaaagctgttttaacaacacaaaggcgattgacttcaca 361
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 217 aaacaaacagcttaaccacaaglytlttcttaacacgacggaatcttcctctta 276
QY 362 gtaacggaactctctatgttcccaacggtgagtcgagggactgagcagggtct 421
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 277 gggaagcggttctctctatcttcttaatttgacaatatattctgctacgttgcaggttct 336
QY 422 gtaacagcagcggtgtagaiaatctacacggttatagggtttcttcgctatctt 481
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 337 gtaacgaatacagcagctctctggtatgtaacgaatctcagattttcaactctcgagtg 396
QY 482 attggtctccctggaagttcgataactacggcaaaaggcggttaagctcctaagggt 541
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 397 ctgacagctccct-----agaccacaggttaaggagccatta--aaattacgagat 444
QY 542 agcttgagttgacaaaatgtcagtttgcctcttcagcaaaaacttttcaacgagtaat 601
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 445 ggtctggtgtttgagaggtataggaatcttgacccaataagaaatgscctctgtgaagt 504
QY 602 ggcggtgtcattacccgcaaaaactcttcaatacagaggaactacaatgtcagctgttt 661
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 505 gggggagccatataatcagaaacttgtcttgcactgaggaagtaagcgttttagagttc 564
QY 662 tctgaaaatacctctcccaagaagaaggcgagcatcagactcagatcccttaccatt 721
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 565 ctgtgcaataagctgttcgacaacaaggggagcgactcagctctgtgactctgtgatt 624
QY 722 actggaacaaagggaagttctcttcttctgacaactcttctcgattctgtgactgca 781
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 625 tctggaatctgaggaatcttgagcttcgaaacaacagcgacgaacatcagagggcg 684
QY 782 atttttaaggaagcctcggtgactatcttcaataatgctcaaatgttccctttatgcaat 841
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 685 atctctgctgaagggaaccttgtgactcctaataccaataatcttcttcagctgc 744
QY 842 aaggtcacagggcgagctccccaacaacaggggagatgtcagggagtgactctgtgt 901
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 745 aaagcaactacaataatggcgagctatgttttaacaagaacagggt----- 790
QY 902 tataaactagtagacagataataggtcaaccctcactggaatcagatgttactctcagc 961
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 791 -----cgaaaccagaccctactcttgcactcttcaggaatgagagctcgatatttcg 843
QY 962 aacataactcgaacacagcgagggaggtactactatgttgaaaaagctcgaaactgccttc 1021
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 844 aataacacagcgaggaatagtgaggtgagatttaccacaaaatattgtgttatcccta 903
QY 1022 ggaagacttacccctatctcagtagaataatgttcaatgaggtgacagctccctaagggtga 1081
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 904 ggaagagagggaggtgttatttcttaacaacaagctcgaaatgctactcctaagaaggg 963
QY 1082 gccatagctatcgaagatgtggggaattgagttatccgcgagatgtgtgacattgttc 1141
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 964 gaatttgcattctagattctgagagatgagatcatcttcgacagatctcgcaatatcatt 1023
QY 1142 ttttaggggaatcagctcctctac-----tactccggagacgaatgaagtgtatc 1195
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1024 ttcgaggggaataactacagcactacaaggaagttccgcgagtgtagccgaatctctaa 1083
QY 1196 gacttaggaacagtgacaaagatgacagcttgcgttctgcgtcgtgtaggacatctac 1255
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1084 gatcttgcagcgaatgcgaatattttaaatctccgcagcagctccggggaataaagtatt 1143
```

```
QY 1256 ttctatgcccataactacagagatcatccacaacagttacagatgtttaaagtatt 1315
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1144 ttctatgaccctatacga-----gtccaaggactactgataaagctctcttggat 1194
QY 1316 ggaactccgagactctgcgactacatacatcaggaacatcatcttccacaggaagaa 1375
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1195 aaagctgcagcaggtctcgaataactcatgaagctcatcatctgttctctcggagaagaa 1254
QY 1376 ttatcagagacagagggcgagatcttaaaaaacttactctgaagctctacagctgta 1435
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1255 ctctcagaagaggaacttaagaaacctgacaaactcgaagcttaccatttaacagctgta 1314
QY 1436 actcttcagagggactctactctttaaacaactggaagtgactctgcagactcagctatc 1495
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1315 gagcttgcctgcaggtgcctcttgaattgaagaatgaggtgacgttattggaatactata 1374
QY 1496 actcaacagcagatctctgcgttcgaaatggaagtagaactactactaga--acttgc 1552
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1375 acgcaggtcgaaggatcgaaagtcgttagtagtagggagactacttttgaggcaagcgt 1434
QY 1553 gatattacacacataaacaattgttcatataacatcagttctatagcaggtgcaagaag 1612
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1435 gagggtgactactcactaagcctagccatataatagattcccttagatgaggacaataaa 1494
QY 1613 gcaaaaatagaaacccaagctacggtcacaataactgacttatacttggaacatcatta 1672
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1495 gctacatataaggagcagcgacgagcaagtaagatgttgcctatccgggscctatcgtct 1554
QY 1673 ttggaacgagcgagcagtttatgaaatacagatttaagaaatcctcagttcctagac 1732
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1555 gtagatgtcagagggaactatattagatcatataatcagtaacagcaggtcttctct 1614
QY 1733 attcagaagctcaaaagctcttcggaactgtaacagacagccaggtgacatccctata 1792
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1615 ttaataagagcttctgcgcaaggaaagatgacttaataagatatcccgatacccaatt 1674
QY 1793 atgggtgagaattccatacagcctacaggaagattggggcccaattggttggggagca 1852
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1675 ctaaatactacgaactacactataggtatcaagaagact---ggaataattgtttggtgac 1731
QY 1853 ggggctct-----agagctcaaacctcacaatggaactaaactgtgctataatcct 1903
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1732 gatgcaactgcaaaaacaaaataatgtctacttaacttgcactaaacaggaatacagcg 1791
QY 1904 aatcccgagctatcggtctctttagctccatagcttaagcttaagatcttatagattt 1963
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1792 aatcagaacgacagggaccttgttctctaataagccctgtgggttcttcttgcagtg 1851
QY 1964 agctctccattatctatgtagagactgcacaacgaaggtgtgcagggagacgtgtctt 2023
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1852 cgctccatccagagctcactatggaacggagcaaggttgcttatctctgtccaacaatttg 1911
QY 2024 tgggtgctggaattactaactcttccataagatagtaglacaaaacagcagcggttt 2083
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1912 tgggtatcagaatcgcgagacttcttgatgaagatcagaaagaaacacacgtatgtat 1971
QY 2084 cgccattgagtgcggtatgttgcataaggaagaaacccaactactgttgcagataagtt 2143
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1972 cgtcattctagcggggtttagcattaggaaggaagattcttcaacggctctcgaaatttc 2031
QY 2144 cttagtctgatttttgcaactcttcttgaaagagataagacactacttctgtagctaaagt 2203
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2032 ttttaatttgccttttttgacgcttttttgctcgaacagagccatctgttgctaaagaa 2091
QY 2204 caaggttaagctcgaaggaggaactctctatcaccagcaacgaacactatatactctt 2263
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2092 cataccatgatatatgcaggggacatgattacgcacaccccgagagtgactaaagacctc 2151
QY 2264 ccttgcaaacacagcgccctgttgcgttctatgttcttcacagagatccctctctctt 2323
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2152 gctaaaga-----tttgcaggaataattctgtactccctcttctgtcttc 2196
```

```

QY 2324 tcaggaacacccatagctacacccatcgataacgatacgtctgaaacccaagtataacacatat 2383
D 2137 aatgctcggttctgcttactgagccatccacataacatgacacaaagatacgtgctat 2256
QY 2384 cctactgttaagaaagctggggaatagatgcttcgttaagaattcggtggagagct 2443
D 2257 tctctgttaagaaagctggggaatagatgcttcgttaagaattcggtggagagctc 2316
QY 2444 ccga---tttcctagatgaagatgcttattttgagcagatcagcccttaagaattg 2500
D 2317 ccggtagttgttcagagcgttcgttggttgatgataccacacgaccttcttaacct 2376
QY 2501 cagttgtctatgacacacagaaaggttttaagaaacgggaacgggaagctcgtaatt 2560
D 2377 gagatgactctgacacacagaaagctttaaagaaacgggaacgggaagctcgtaatt 2436
QY 2551 ggaagtagcgtctgtgtaattcttgccttaacatcagcagacgagcttgaagaatca 2620
D 2437 caaagtgaagaccccttcacatcagcttcgttgagtgataaatttgagaa-----a 2490
QY 2621 gactgccaagatgcaacgtataacatcttgggtatctgtgagatctgttcgtagt 2680
D 2491 ttcccgtaagatgcaacgtataacatcttgccttaacatcagcagatgagctaat 2550
QY 2681 aaccgcagctgacacacacacacacacacacacacacacacacacacacacacacacac 2740
D 2551 gacacagctgacacacacacacacacacacacacacacacacacacacacacacacac 2610
QY 2741 aattggaacaaagacacacacacacacacacacacacacacacacacacacacacacac 2800
D 2611 agctgtgtaacaaagacacacacacacacacacacacacacacacacacacacacacac 2670
QY 2801 ttgaaacacacacacacacacacacacacacacacacacacacacacacacacacacac 2860
D 2671 ttgaaacacacacacacacacacacacacacacacacacacacacacacacacacacac 2730
QY 2861 gacttagaagcaaaatcccaattctaa 2887
D 2731 gacttagaagcaaaatcccaattctaa 2757

```

```

RESULT 7
X06822 standard; DNA; 2787 BP.
AC X06822;
DT 26-APR-1999 (first entry)
DE Chlamydia pneumoniae surface exposed protein Omp10 DNA.
KM Omp10; outer membrane protein 10; surface exposed protein; antigen;
KW infection; diagnosis; vaccine; atherosclerosis; asthma; ss.
XX Chlamydia pneumoniae.
OS Chlamydia pneumoniae.
PN WO9858953-A2.
PD 30-DEC-1998.
PF 19-JUN-1998; 98WO-DK00266.
PR 23-JUN-1997; 97DK-0000744.
XX (BIRK/) BIRKELUND S.
PA (CHRI/) CHRISTIANSEN G.
PI Birkelund S, Christiansen G, Knudsen K, Madsen A;
PI Mygind P;
DR WPI: 1999-105610/09.
DR P-PSDB; W88423.
XX

```

```

PI Species-specific test for identifying mammals infected with
PI Chlamydia pneumoniae - comprises detecting antibodies specific for
PI outer membrane proteins of C. pneumoniae or nucleic acids encoding
PI these proteins
PS Claim 6; Page 59; 115pp; English.
XX
CC This DNA sequence codes for the novel 98.4 kDa surface exposed
CC protein Omp10 (see W88423) of the human respiratory pathogen
CC Chlamydia pneumoniae. By generating antibodies against C.
CC pneumoniae outer membrane complex, a polyclonal antibody (PAB 150)
CC was obtained which reacted with outer membrane proteins. The
CC antibody was used to identify the genes (see X06816-27) encoding
CC Omp4-Omp15 proteins (see W88417-28) in an expression library of
CC C. pneumoniae DNA. The genes are situated in 2 gene clusters:
CC Omp12,11,10,5,4,13 and 14 in one cluster and Omp6,7,8,9 and 15 in
CC the other, and encode polypeptides of about 89.6-100.3 kDa and
CC about 56.1 kDa. The invention provides a new species specific test
CC for identifying mammals (including humans) infected with Chlamydia
CC pneumoniae. The test comprises detecting antibodies specific for
CC Omp4-Omp15 or detecting nucleic acid fragments encoding these outer
CC membrane proteins, especially by PCR. The proteins are also used
CC in the diagnosis of C. pneumoniae infection in mammals. The
CC nucleic acids and proteins can also be used in the immunization of
CC mammals, the nucleic acids being particularly useful as DNA
CC vaccines for effecting in vivo expression of antigens. The
CC vaccines may also prevent atherosclerosis and bronchial asthma,
CC which are possibly associated with C. pneumoniae.
SO Sequence 2787 BP; 815 A; 689 C; 535 G; 748 T; 0 other;

```

```

Query Match 17.2%; Score 515; DB 20; Length 2787;
Best Local Similarity 53.3%; Pred. No. 1.3e-132;
Matches 1425; Conservative 0; Mismatches 1175; Indels 75; Gaps 13;

QY 243 aagatcgtgaactactactactttaaaggaaatgctactctagaataattctgtgaa 302
D 158 atgcaatgtgaactactactactactacacagggatgtctc---aacacacacagcagat 214
QY 303 caggcacagaacacacacacacacacacacacacacacacacacacacacacacacacac 362
D 215 ctccgacagctcacaacgcgtctctgtttaaagaacacacacacacacacacacacacac 274
QY 363 gtaacgggaactctctattgttccaaacggtgagtcaggaactgtagaagggtctgtg 422
D 275 gccacggctacaacttctccctacaacacacacacacacacacacacacacacacacac 328
QY 423 ttacagcgagcgtgtagtaatactcaccagtttaaggtttcttcgcacatttta 482
D 323 ttacacatacagctgaataaagctctctcttccagatctctctatttgcactaa 388
QY 483 ttgcgtctcctggaagttagatactaccgcaaggaacggttagctgtctacgggta 542
D 389 taacaacacagatgtc-----accacaggaacagggacccaagtcacagaagctt 442
QY 543 gcttggtttgacaaaatgtcagttgtcttccagcaaaaacttttcaacggaatg 602
D 443 gtctctatcagtcgaactat---agttgtcatttggccaaaacttttcaatgacaatg 499
QY 603 gcggtgtataccacgcaaaaactcttcaatacaggggacacaaatgtcagtcgtttt 662
D 500 gagggcctctccaaagcagctctatcagctctcagctcaaacccaacctaac---gttg 556
QY 663 ctgaataatcctctccaaagaaagcgggacatccagactccgatgccttaccatata 722
D 557 ccaaaacaaagcaagcgaagaaaggggtgcctctattccacgggaggtatcaatata 616
QY 723 ctggaacccaaggggaagctctttttctgacaactcttcgagatctgagagctgaa 782
D 617 acaatcgtttaactcagatcatcttctgaaataacacgcgcggaacaaatgcygagca 676
QY 783 tttaacagaagcctcgtgactatttccataaataagcgaagtttccttattgacata 842

```



DB 677 ttacacggaagctagcagtttatttagcgcaacaagaacattagctttataacata 736  
QY 843 aggtcacaggagcagctcctcaacaacggggaatgtaagggggtcctatcgctc 902  
DB 737 gtgtgacg-----caacctcagctacagggggccattaacgtgta 781  
QY 903 ataaactagctacagatacctaaagctccctcactgtgaaatcattgtactctcagca 962  
DB 782 gtacatcagcccccacaacacgcttactctatcagcaacggggaacttag 841  
QY 963 acataatcgcacaacacggggaagctatctatgtgaaagctcgaactgctccg 1022  
DB 842 gaataaagaatattcagttgtgtggggttactgtcaactatgtcttctctg 901  
QY 1023 gagactacccctatctcagtagaatagttcaatgtgaggtacagccctcaaggtgag 1082  
DB 902 gaggaactaagcttttaaaaaaacacctctatagatacgcagcccttaaggagag 961  
QY 1083 ccataagctacgaagatagtggggaattgagttatccgcgagtagtgtgacatgtct 1142  
DB 962 caattgagctatgctgactcgtgactcttgagcttctcgctcgtgtgtggagactcct 1021  
QY 1143 ttttaggaatacgt-----cactctactactcctctggagcgaatagaagtga 1193  
DB 1022 ttgaaggaaacacagtgtaaaaggagcttctcagctcagacacattacgaatattcta 1081  
QY 1194 tcgacttaga--acgagtgcacaagatgacagcttgctcgtctcgtcgtgtagagcca 1250  
DB 1082 ttacatcgggaacacacaaatgtaagattgtacagctcgtcgagcctccaaggcaattcta 1141  
QY 1251 tctactctatgatacctataactaagatcattccacaaagttacagatgtcttaaaag 1310  
DB 1142 tctactcttctgtcctatacaactaactaactgagctctctcagatcgtcttaact 1201  
QY 1311 ttataggaatccgcgacgattctgactatacaatatacagggaactattcttcacagag 1370  
DB 1202 taatagctcgtgactcgtcggaagatcctcattacaaagacacatgcttcttctgag 1261  
QY 1371 aaaaatatacagagacagggcgagatctcaaaaatcttactctgaacatacagc 1430  
DB 1262 agagctctcgcgaagcagagctcgaagctgtaattcccaattcaacttgaagac 1321  
QY 1431 ctgtaatcttccaggggtactcttacttaaacatggaagctcctcgagctag 1490  
DB 1322 ccttaactctggggggggaactctcttaaatcagagtaactctagtgtgtaagt 1381  
QY 1491 cattcaactcaacagcagattctcgtctgaaatggaagctgaggaactctagaactg 1550  
DB 1382 ccttttcgcaatctcgggctcctacccctcctcattgagtcaggggaacattagaaccg 1441  
QY 1551 ctatatactagcaccaataaacttggctatcaatcagttctatagaaggtgcaaga 1610  
DB 1442 ctgagggatcacactcaataatctgtctcattgagttgagttccttaagaagaccca 1501  
QY 1611 aggcacaaatagaacccaagctacgtcacaanaatcgtacttattctgggaacatact 1670  
DB 1502 agactaagctaaagacacgaagcaagtaagacagtaacttattctggaatcgtctc 1561  
QY 1671 tatggacccgagcagcttattatgaaatcagtttaagaatccctcagctcag 1730  
DB 1562 ttgtagactccttggaatgctcaggaagatgtctcttgaaataccctcaagctctt 1621  
QY 1731 acctctagagctcaagct-----ctggaactgtacaagacgcgagtaactcag 1784  
DB 1622 ctgtctcaacttctacgtcagaccccgaaatattcacatcagtaactgctgctg 1681  
QY 1785 atccataaaggtgagaattccattacggtatcaggaacttggggccaatgttt 1844  
DB 1682 atccctcagaanaaaatccctatccattgggataccaaggaattggtcattcctggc 1741  
QY 1845 gggggaacagggc-----ttctaagactgcaaccttcaactgactaataacgtgctata 1898

DB 1742 aagagatactcgcgaactaaatcccaagcagcgaactcttactctgacaaaacagagata 1801  
QY 1899 ttccatcccgagcgatagcgtctttagtccctatatagcttaatgaaatgattatag 1958  
DB 1802 atccgaatccctgagcgtcgtggaaccttagtctacaacgctatggggaatctctgtg 1861  
QY 1959 atattgctctcccttattctatgtagagactgcaacggaaggtctgcaaggagacgtg 2018  
DB 1862 atgtgctctcatalacaacgctgtgagccactaaatagcgaatctcaagaacctcg 1921  
QY 2019 ctttgggtgtgattatcctaactctctcacaagaagtagtaaaaaacagacg 2078  
DB 1922 gcattcgtgtgaaaggatctcgaactctctcacaagaagtagcagaataataaag 1981  
QY 2079 ggtttgcacattgagtggtggtatgtcacaagaagaaacctaactactgtctgata 2138  
DB 1982 gtttccgcacaaatgacggttattgttaggaagcagatacaacattagctctgata 2041  
QY 2139 agattcttagtctgcatattgtcagctcttggaaagataagactactgtatgata 2198  
DB 2042 atcttactagcagcctctcgtcaattattcgggaagatagagttccttataata 2101  
QY 2199 agaatcaaggtacagctcagaggaactcctattacaacgacaagaaactatact 2258  
DB 2102 aaaaatagactctcgtcagctcctctccatctccagcactagcagactgtctt 2161  
QY 2259 cctctcctgcaacacagcgcctgtctgtgtctatgttctctacagagatccgcttc 2318  
DB 2162 ctccaagctgttac-----gctacccctcgtgactcgaagtgagcagcgtctcc 2212  
QY 2319 tctttcaggaacacttagctacacacacacagataacgatacgtctgaanaacaglataca 2378  
DB 2213 tctttgagctcagatcagctatatactatgaaataactactaagaacactatacc 2272  
QY 2379 catatcctactgtttaagaagagctggggaatgatagttctgcttgaagattcg--gtg 2435  
DB 2273 aagcacacaaaggagagagctcgtgtataatgacggtctcgtcgtggaactgagcgt 2332  
QY 2436 gaagagctccgacttcttagatgaagagctcattatgacagctacagcctctatga 2495  
DB 2333 cctctacacacactgtcttaagcagaggtctctcctcagcggtatttctcttataca 2392  
QY 2496 aattgagttgtctctgacatcaggaaggtttaagaac--aggaacagaaagctc 2552  
DB 2393 aagtagaagctctgatacacaacaaagatagcttcaagaacgtaatactactgtgac 2452  
QY 2553 gtgaatttgaagtagcgtctgtgtaactctgcttactatcaggatccgatttgata 2612  
DB 2453 gatcttctgataagcgttgatttaataacgctcgtgcttattggaattacattcga 2512  
QY 2613 aggaatcagactcccaagatgcaagctacatcctaactctgtgtatgtgagctg 2672  
DB 2513 gatctcgggaagacgagcgtcgtcttaagaagctactctcactcgttgcgagatgct 2572  
QY 2673 ttctgttaaccccgacttcaagacaacactcggaaattgagcgtgtattcttggaaact 2732  
DB 2573 atcgttaagaatccctgactcagcagacgctccttaatacaacataactcgttga 2632  
QY 2733 tcggtacgaatttggcaaaagacgcttagtctctgtcagaggaacacttttcttata 2792  
DB 2633 caggaagaaatctctcaagaacagcgtgatacggaaagacaggaactttaaagctct 2692  
QY 2793 actcaaatttgaagcctttagccaatttcttcttgaattgctggtgcatcgcgaact 2852  
DB 2693 ctccaactcttggtgcacaagtaacctaataatgaaattcgttgattcttaacgagct 2752  
QY 2853 acatgttagacttaggaagcaataatacaacttctaa 2887  
DB 2753 acaatgagatcttggaggtaagttccagttctaa 2787

RESULT 8  
X06823

ID	XX	X06823	standard; DNA; 2793 BP.
XX			
AC	XX	X06823;	
DT	XX	26-APR-1999	(first entry)
XX			
DE	XX	Chlamydia pneumoniae surface exposed protein Omp11 DNA.	
KW	XX	Omp11: outer membrane protein 11; surface exposed protein; antigen;	
KW	XX	infection; diagnosis; vaccine; atherosclerosis; asthma; ss.	
OS	XX	Chlamydia pneumoniae.	
XX			
PN	XX	W09858953-A2.	
XX			
PD	XX	30-DEC-1998.	
PF	XX	19-JUN-1998; 98WO-DK0266.	
PR	XX	23-JUN-1997; 97DK-0000744.	
XX			
PA	XX	(BIRK/) BIRKELUND S.	
XX		(CHRT/) CHRISTIANSEN G.	
PI	XX	Birkelund S, Christiansen G, Knudsen K, Madsen A;	
PI	XX	Mygind P;	
DR	XX	WPI: 1999-105610/09.	
XX		P-PSDB; W88424.	
PT	XX	Species-specific test for identifying mammals infected with	
PT	XX	Chlamydia pneumoniae - comprises detecting antibodies specific for	
PT	XX	outer membrane proteins of C. pneumoniae or nucleic acids encoding	
XX		these proteins	
XX			
PS	XX	Claim 6; Page 62-63; 115pp; English.	
XX			
CC	XX	This DNA sequence codes for the novel 97.6 kDa surface exposed	
CC	XX	protein Omp11 (see W88424) of the human respiratory pathogen	
CC	XX	Chlamydia pneumoniae. By generating antibodies against C	
CC	XX	pneumoniae outer membrane complex, a polyclonal antibody (PAB 150)	
CC	XX	was obtained which reacted with outer membrane proteins. The	
CC	XX	antibody was used to identify the genes (see X06816-27) encoding	
CC	XX	Omp4-Omp15 proteins (see W88417-28) in an expression library of	
CC	XX	C. pneumoniae DNA. The genes are situated in 2 gene clusters:	
CC	XX	Omp12,11,10,5,4,13 and 14 in one cluster and Omp6,7,8,9 and 15 in	
CC	XX	the other, and encode polypeptides of about 89.6-100.3 kDa and	
CC	XX	about 56.1 kDa. The invention provides a new species specific test	
CC	XX	for identifying mammals (including humans) infected with Chlamydia	
CC	XX	pneumoniae. The test comprises detecting antibodies specific for	
CC	XX	Omp4-Omp15 or detecting nucleic acid fragments encoding these outer	
CC	XX	membrane proteins, especially by PCR. The proteins are also used	
CC	XX	in the diagnosis of C. pneumoniae infection in mammals. The	
CC	XX	nucleic acids and proteins can also be used in the immunization of	
CC	XX	mammals, the nucleic acids being particularly useful as DNA	
CC	XX	vaccines for effecting in vivo expression of antigens. The	
CC	XX	vaccines may also prevent atherosclerosis and bronchial asthma,	
CC	XX	which are possibly associated with C. pneumoniae.	
XX			
SQ		Sequence 2793 BP; 803 A; 681 C; 567 G; 742 T; 0 other;	
Query Match	17.1%;	Score 514;	DB 20; Length 2793;
Best Local Similarity	52.7%;	Pred. No. 2,4e-132;	
Matches 1404;	Conservative 0;	Mismatches 1200;	Indels 60; Gaps 11;
OY	250	tggaactactcatttaagggaatgtcactcgtacagaatatcttcggagacagc	309
Db	162	tggaacgaactatgcttatacagaatgtctatat--aacgatgtcgggaagagcc	218
OY	310	agcaatcacaanaagctgtttaacaacactaaggcgagtttgactttacaagtaag	369
Db	219	agcatataacaggctgcgttatacagaacactcgggtgtatcttcgaattacgtgaag	278

QY	370	gaactctcatatggtcccaacggtgtatgacgaggaactgtaaaggggtcgtctaaacag	429
Db	279	atactatttttaattcaacaacggtatgacgtgtctgaatctcaaggaaactgcgg---caag	3353
QY	430	caagctgtgtagtaataactacacacagtttatatgggtttttctctgcgtacttttatgttc	489
Db	336	caacaactgtctgataagccctaacaacttcaacaggaatttttctaactcttccttaacagc	3953
QY	490	tctctgaagtctcgataaactacgcggcaaaaggacggttagctgtctacggtatgtgag	549
Db	396	tctctgaactaaacagtgtcttcagagaaaaagtaacttaa---gtctcagagaccttaa	452
QY	550	tttgcgaaaaaatgttcaggtttgcctcttcagaanaaacttttcaacg-----ataa	600
Db	453	tcttaacgtaataatgtgaacagatctctcttaagcnaaacgctcccaatgaactaatacaa	512
QY	601	tggcggtgtatcaacgcgcnaaaactcttcaattcaacagggactcaaatgtcagctgtt	660
Db	513	tggcgaggtatcaacacacnaaaactcttctattcttctgggaataacctctctaact	572
QY	661	ttcttgaanaatacctctctcaagaagaaggcgccattcagaacttcgactccattccat	720
Db	573	caactgtataacgcgaaaaaaattagtggtgacgtataagctctcgtcgttcaagttat	632
QY	721	tactcgnaaaacccaggggagagctctctttctgcgaacaactcttcgactatctgtgaatg	780
Db	633	ttcaggaanaacacgcgcagcttagctcttatgtataataaagggaacactggtggcggtg	692
QY	781	aatttttaacagaaagctctgcgtgactatttcttaataatgctcaaaagttctcttatgt	840
Db	693	tctgtggtcttgaagacagctccctcgtatctcaaaaatagctccctcttctctctgaa	752
QY	841	taaggctcacagagacgagctcctcaacaacggtgatatgtcagggagtctactcgtgc	900
Db	753	caactgcacagatgctgagggcaag-----ggggggccacttatgt	794
QY	901	ttataaactgtgtacagataactaaagtccacccctcaactgtgaatcagaatgttaacttcag	960
Db	795	tgaaaaaacagagagagactcctactcttaactctcctgcgaataataaaagtctgacctgcg	854
QY	961	caacatacacgcgcagacagcgggagagctactatgctgtgaanaagctcgaaactgtcttc	1021
Db	855	cgaagacctcttcagtaactcaaggcggagaaactgtgcctacgtctcgtacatcttcgcgc	914
QY	1021	cggaggaacttccctattcagtagaataagtgfcaatlygaggtacagcttccctaagtttg	1080
Db	915	tgtctgcctcctcctatttttaaaatagatgcggaacaacagctgcaggaaggcgcg	974
QY	1081	agccatacctatcgaaagatagtggtggaaatgtgattatccgcgcgaatagtggtgaaattgt	1144
Db	975	cgcatatgcatttgcgcacccctgcgaatctttaagctctctcgcgaacaagagggaactcac	1034
QY	1141	cttttttagggaatacagltcaacttta---ctactcttggagacatagaaagtatgatacga	1197
Db	1035	gttctcttgcaaacctcttaaccttaacctccgcgcgaacatgcgaacggaaatgcatctta	1094
QY	1198	ctttaggaacgggtgtcaaaagtgtacagtttggtctgtcgtctgtatagagccatctactt	1255
Db	1095	ctgtgtgatacgcacgcaaaaatttcgaacttaagggtgagcccaaggccaatcttatattt	1154
QY	1258	ctatgtatccataactacagatcatccacacaaacagttacagatgtcttaagaagttaatga	1311
Db	1155	ctatgtatcggat---tgcatactaacacaaagaggtcttagaggttcttgcacatcaacca	1211
QY	1318	gaccgcgcaggaattctgtcactataacatacaggggaacatcatcttcaacaggaanaagt	1377
Db	1212	acgggatagcacaaccgcctttagatattatcagaatgaagatgttatcttctgtggaaaaagt	1271
QY	1378	atcagagacagagccgcgcaggaattctaaaaactcttacttggagagctactacagctgttac	1437
Db	1272	ctctgcagatgtgaagcgaagctgtctgataacttcaacatctatataataaagcaaccattg	1331

QY 1438 tctttcaggaggtactctatctttaaacatgagtgactctgcagactcagcgtatcac 1497  
 DB 1332 tctagcctcttggaaccttgacactcaagaatgtcgtgtagtgcataatggtttcc 1391  
 QY 1498 tcaaacaggagatctctgtctcgaatgtgacgttagaactactct--agaacctgt 1554  
 DB 1392 aagaagctgaggtcttaccctctccatgaaccaggaacaaagctcaagcagatctga 1451  
 QY 1555 tactagcacataaacatgttgctatctacatcagttctacagcgtgcaagaagc 1614  
 DB 1452 agctatcagctcttaccacactgtgtgtatcttctcgtcttagaggaataagagtg 1511  
 QY 1615 aaaaatagaacccaagctacgtlcaaaaaatctgacttattctggaaccatctatt 1674  
 DB 1512 gtcacttgaaacagcagcgagccacaaactataacttaactctctcctgtttcca 1571  
 QY 1675 ggaaccgcagcgacagtttatgaaatcatatgtttaagaatctctgaactcgaact 1734  
 DB 1572 agatgtagcggaattttatgaaagccatagaaacccaagccttaccgagcctt 1631  
 QY 1735 cttagagctcaagcttctggaactg-----taacagcacccgagtgactccaga 1785  
 DB 1632 ggtgttatctacactgtctgactgctgctagcgtatattatcagatgctctcacttc 1691  
 QY 1786 tccataatgggtgagaaatcccatcagcgtatcaggaagactggtggcccaattgttg 1845  
 DB 1692 tccagttacaaactcagaacccttaccggtatcaggaactgtggaagcacttgagc 1751  
 QY 1846 ggggacaggggttctacagctcgaactcgaactgaactgaactgaactatctccta 1905  
 DB 1752 aagacatacaactcgaacaaatcaggaactatgacttggtaactacgtgtaaacctaa 1811  
 QY 1906 tcccgagctacggtctctttagctccatagctgtaatgtaactatataatag 1965  
 DB 1812 tccgtgagctagagcttccgttagtcccgatcattatgagctccttaccgacatttg 1871  
 QY 1966 cctctccatctcttaagagagcgaacgaaggttgcaaggagacgtgtcttttg 2025  
 DB 1872 cactctacagcagatcatatcctcaagcagatagctatcagcaagcagagctctg 1931  
 QY 2026 gttgtgattatcttaactcttccatagagtagtataaacaacgagcggttttg 2085  
 DB 1932 ggcacacaggaactcgtgaattcttccatagagataatcaggaactcaacgaactcgg 1991  
 QY 2086 ccaattgagtggtggtatgtaacaggaaggaactcactatctgttcagataagattct 2145  
 DB 1992 acataaagctacggtcatatgttggaagagtgctggaagatttctcgaataatctt 2051  
 QY 2146 tagtctgcatcttgtagctcttgggaagagatagagctactctgttagcaaaatca 2205  
 DB 2052 cagtgtagctctctgcagctctcgtgtaagataaagacgtgttatgttgaataatc 2111  
 QY 2206 aggtacagltcagaggaactctctatcaccagacgaacgaactatctctcttc 2265  
 DB 2112 ctctcataactatttagcgtcgtctatcactgcaacatcagatctccatagaggtcttc 2171  
 QY 2266 ttgcaactagcgcctgtgtctgttctatgttctctacagagattcctgtctcttc 2325  
 DB 2172 --catgcccatttggaaatcatcaccgacatgtgtaagatattcctcattttgaa 2228  
 QY 2236 aggaaccttagctacaccatcagatagatcgtgaaacaaagttatacaactatcc 2385  
 DB 2229 tgcacagctaaagctacacatacaaaaatgatatgatactcgtctataactctatcc 2288  
 QY 2386 tactgttaaggaagctggggagatgattcgtcttaagattcgttgggaa---gagc 2442  
 DB 2289 tgaagctaaagttcttggaaccaataatctcgtgggtctcagagcgtgggggtctcgtgc 2348  
 QY 2443 tccgattggtctatgtagaagagctctatcttagcagcatcagccctcactatgaaattga 2502  
 DB 2349 tctatactccctaaagacacgcttctccaggagatacttcccaattttaaaattcca 2408  
 QY 2503 gttgtctatgacatacaggaaggtttaagaacaggaagaaagacgagctggaatttg 2562

DB 2409 ggcagttcaacgcccgcacaacaaacttcaagaagagtggtcgtgagccgtgtttga 2468  
 QY 2563 aagtagccgtctgttgaattctgtccttaccatagacgtggatccgatttgaagaatcaga 2622  
 DB 2469 tgaatgaagccttagagactgctcatccctcgtcgtcgtcgtttagaanaaatccga 2528  
 QY 2623 ctgcaagaatgaaagctacatctcaactcgttgatctgtgtatctgtgtcgtgagtaa 2682  
 DB 2529 agatgaaaaaaataattcgaattctccttagccaacttggatgtgtagtcaaaaaa 2588  
 QY 2683 cccgcactgtacgaacaacactgagatagcgtgtatctcttgaanaacttcgtgacaa 2742  
 DB 2589 tcccggttcggtacttctcgaatgtcagtgagtgagccctcgtgacttgcatagtaaaa 2648  
 QY 2743 ttgtagaagacagctttagtctcgtcgtgcaagggaacaaatttgccttaactcaatt 2802  
 DB 2649 cctcgcagacagacgctctctttagcaagtgctggaagccatctgactctccctcatgt 2708  
 QY 2803 tgaagcctttagccaatttcttcttgaattcgtgtgcatctgcgaattcaaatgtaga 2862  
 DB 2709 agaactctcgtgggaagcgtcttatgagctcgtgtgctcagacacacatacaatgtaga 2768  
 QY 2863 cttagagcaaaaataccaattcta 2886  
 DB 2769 ttgtggttaagatactatctcta 2792

RESULT 9  
 X06828  
 ID X06828 standard; DNA; 3000 BP.  
 XX  
 AC X06828;  
 XX  
 DT 26-APR-1999 (first entry)  
 XX  
 DE Chlamydia pneumoniae surface exposed protein Omps DNA.  
 XX  
 KW Omps; outer membrane protein 5; surface exposed protein; antigen;  
 KW infection; diagnosis; vaccine; atherosclerosis; asthma; ss.  
 XX  
 OS Chlamydia pneumoniae.  
 XX  
 FT CDS  
 FT CDS 259..3000  
 FT CDS /\*tag= a  
 XX  
 PN W09858953-A2.  
 XX  
 PD 30-DEC-1998.  
 XX  
 PF 19-JUN-1998; 98WO-DK00266.  
 PF  
 PR 23-JUN-1997; 97DK-0000744.  
 PR  
 XX  
 PA (BIRK/) BIRKELUND S.  
 PA (CHRISTIANSEN S.  
 PI Birkelund S, Christiansen G, Knudsen K, Madsen A;  
 PI Mylind P;  
 DR WPI: 1999-105610/09.  
 DR P-PSDB: W88429.  
 XX  
 PT Species-specific test for identifying mammals infected with  
 PT Chlamydia pneumoniae - comprises detecting antibodies specific for  
 PT outer membrane proteins of C. pneumoniae or nucleic acids encoding  
 PT these proteins  
 XX  
 PS Disclosure: Page 73-77, 115pp; English.  
 CC  
 CC This DNA sequence encodes the novel surface exposed protein Omps  
 CC (see W88429) of Chlamydia pneumoniae, a human respiratory pathogen.

CC It is described as a subsequence of a claimed nucleic acid fragment  
CC (see X08817) encoding Omp5 (see W88418). The invention provides a  
CC new species specific test for identifying mammals (including  
CC humans) infected with *C. pneumoniae*. The test comprises detecting  
CC antibodies specific for surface exposed proteins Omp4-Omp5 (see  
CC W88417-28) or detecting for surface acid fragments encoding them (see  
CC X06816-27), especially by PCR. The proteins are also used in the  
CC diagnosis of *C. pneumoniae* infection in mammals. The nucleic acids  
CC and proteins can also be used in the immunization of mammals, the  
CC nucleic acids being particularly useful as DNA vaccines for  
CC effecting in vivo expression of antigens. The vaccines may also  
CC prevent atherosclerosis and bronchial asthma, which are possibly  
CC associated with *C. pneumoniae*.

XX Sequence 3000 BP; 867 A; 597 C; 658 G; 878 T; 0 other;

Query Match 16.9%; Score 506; DB 20; Length 3000;  
Best Local Samlarity 52.6%; Pred. No. 4,1e-130;

Matches 1383; Conservative 0; Mismatches 1170; Indels 75; Gaps 10;

QY 247 tgcgtgaactactctactttaaagggaatgtcacctagaanaatattcctgacagc 306  
DB 414 tactggaataagactactactctgacagagataataactctgcacaacctgggattcg- 472  
QY 307 cacagcaatcacaaaagctgttttaacaacacgtagagcgacttgacttccacagtaa 366  
DB 473 --cagcttaacgaagggttcttctctgacactcgaacttttaagcttgcgcgta 530  
QY 367 cgggaactctctatgttccaaacggtgagtcaggaactgaagagggctgcgttaa 426  
DB 531 ggggactacacttctcttttaataatagctagctgctgaaggcgagacttctcgt 590  
QY 427 cagcagcgtgtgtagtaatactacacgcttataaggttcttcgtactattatgac 486  
DB 531 tacaactgataaatactcgtcgtacacaggaattcgaagtcttacttctttagcgccc 650  
QY 487 gtctcctggaagttcgataactacgcgcgaagcgccctgcgcgtctaaaggtagctt 546  
DB 651 atcactcgtaatacacaacccctcgaagaaaggtgcagttaatgtggaagggtc---t 707  
QY 547 gagtttgcaaaaatgtcagttgtctcttcaagaaaacttttcaacggaataatggcg 606  
DB 708 tacattgtatacaatggaactatttattttaaacaagattactgtgagaaatggcg 767  
QY 607 tgcataccccaagaactcttcatataacagggagactacaatgtcagctcgttttctga 666  
DB 768 agccatttctacccaagaactcttcttgaaaaacaacagcagggatcgaattctttgaag 827  
QY 667 aaata-----ctctcacaagaagagcgagccatcagaactcgcagtccttac 717  
DB 828 gaataaatacgcgcgaacaggaagaaagtggtgcatctgtactcgtactgtaga 887  
QY 718 catcactggaacccaagggtgaggtctcttttctgaaataactctcgtgactctgagc 777  
DB 888 tattacaataactcgcgtcctcactcctcctcgaacaataatgtcgtgaagctgag 947  
QY 778 tgcgaattttacaggaactcgtgtgactattcttaataaactgaagtttcccttattga 837  
DB 948 agctataaataagcagaggaactgtataactacaggaataagctcctgttatttctga 1007  
QY 838 cataaagttacacggagcggtcctcctaacaacgggggataatgacaggtgatactg 897  
DB 1008 aaatagtgtagc---agcgacgcgcaggaatgtagggagctcttcgagagatgcg---- 1060  
QY 898 tgcattaaactagtaacagatacagtaagtcacccctcaggaactcagatgttactctt 957  
DB 1061 -----atgttaccataatctgggaatacagaatgttaacttt 1094  
QY 958 cagcaacaatatacgcacaacagcgaggagagctatcatgtgaaaaagctcgaaatggc 1017  
DB 1095 ctcaaggaacacagcgtgtagtaatgtagcgagccatttatgtcctaagaagacttaactg 1154

QY 1018 ttccggagga-----cttaccattcagtagaataatgtaacatgaggtacagctcc 1071  
DB 1155 ttccgggggggggggggttactcttcttcttcaaaaatagatccaaaggtaccctgaag 1214  
QY 1072 taaagttgagccataagctacacgaagaatgtaggggaattgagttacccgcgatatgg 1131  
DB 1215 taatggtgagccattctctatcactcgcagctgtagaggtgtagtcccttcaacaagaagc 1274  
QY 1132 tgaactgtctcttttaagggaatacagctactactactc---tggagcgaatagaag 1188  
DB 1275 ggaacttacccttcaatggtgagtgcatgttgtaactacacccaactcaaaaagaa 1334  
QY 1189 tagtatcacttaaggaacgcgtgcagaagtagaagccttgcgtctgcgtgtgtagac 1248  
DB 1335 ttctatgacataggaactctcgcgaagaacacgaatlttactgcaatatcttggcatag 1394  
QY 1249 catctactctatgataccataactacagagatcatccacaacagttacagatgtctaa 1308  
DB 1395 catcttcttccatcagatcgcgtactactgctaaatagcgtcgcgattccacagatactaa 1454  
QY 1309 agttaatgagactccgcgcgaatctcgcacataatatacaggaacatacttccaaag 1368  
DB 1455 tccaataaagcgcgtagcaggtatagtaacagatlaagtggtcgatgttttctg 1514  
QY 1369 agaaaagtatcagagacagagccgcagattctaaatcttacttgcgaactacta 1428  
DB 1515 tgaagaagctctcgaagatgtaagcaaaagttagagcaaccccaactcctcagctgaaga 1574  
QY 1429 gctctgtactcttccagaggtactctatctttaaacaatgagttactctgacactca 1488  
DB 1575 gctgttaactcttactcgcaggaatttagtacttaaacgtgtgtcactcgtatagca 1634  
QY 1489 ggcattcacccaacagcgagatctcgtctcgaataatgagcgtgaagaaactcctgaacc 1548  
DB 1635 aggcttactcagacccggcggtctcctcgttatttaagtggtgggcacaaactttaaagc 1694  
QY 1549 ---tgctgatactagacataaacaattgtgtcaatacagattcttatagagctgac 1605  
DB 1695 aagtaagaggaagtcacttaacaggtcttccacttctcgtgagactctttagaggaagg 1754  
QY 1606 aagaagagcaaaaatagaacacaagctcagtcacaaaactcgtacttatctggaacat 1665  
DB 1755 taagaaagttgttaattgtcgtctcgcagcaagttgaataatgagcccttaagtggtcc 1814  
QY 1666 caattatggaacccgcagcgacgttttatagaataatagtttaagaataatcccgctc 1725  
DB 1815 tctcttcttgataacccaagggaatgacttaagaataatcagacttagaanaaaccaaga 1874  
QY 1726 ctaagcaatcttaagctcaaaagctcttggaaactgtaaacagcagcgatgactcaga 1785  
DB 1875 ctttcaattgtgcagctctcgtcctgtgactcgaacaactacagatgttccagcggt 1934  
QY 1786 tccataaatgggtggaatcattcattacaggtatagaaggaacttggggcccaattgtttg 1845  
DB 1935 tccataagtagaactcctcagcactataggtatccaagaatgactcgtgggaatgagttggt 1994  
QY 1846 gg-----ggacaggggtcttcaacgactgcaacactcgaactggaataaacgg 1893  
DB 1995 tgatgataccgcaagcactccaagaactaagaacagcgacatagcttggacaataacag 2054  
QY 1894 ctatatctcctaaccgcgggtatcgtctctttagtccctaataagcttatggaatgact 1953  
DB 2055 ctacttccgaatcctcgcagcgtlcaaggacctttagtctcctaataagccttgggatactt 2114  
QY 1954 tataataatgctctcccaattatctatgagaaactgcaaacgaaggtgtgcagggaga 2033  
DB 2115 ttcagacatccaagagatccaaggtgtcatagagaagaagtgcttgaactctgtttcaga 2174  
QY 2014 ccgtctcttctggtgtcgtgatacttaactcttccataaagatgatacaaaacag 2073  
DB 2175 tcaaggtcttctggtcgtggtgagtcgccaattctttagataaagataagaaggggagaa 2234  
QY 2074 acgcggttccgcaattgagtggtgttatgtcatagaggaacactacataactgttc 2133

Db	2235	acgcaaataccgcataaactctgttgatagctatcggaggttgacgcgcaaacctgttc	2294
Oy	2134	agataagatctcttgctgcctgcattttgtctacgctcctttggaagaagatagactattgt	21939
Db	2295	tgaanaacttaattgaccttgccttttgcacacaccttttggtaagatlaagaatttcttagt	2354
Oy	2194	agctcaaatcaagaagtgacagctctacgaggaactctctctgtccagcaacaagaaacta	2253
Db	2355	cgcataaataatcatctgtatcatccatgcaggaagcctctataccaaacatatcaagatg	2414
Oy	2254	tatctctctctctgcgaactacagcgccctgttctgtgtcttaatgtctctacagaagatcc	2313
Db	2415	tagtgggtttcataagttgctctctctgaataaactctctgctcttggatgcataaac	2471
Oy	2314	tgctctctcttcagaagaaccttagtaaccccaataagataagatcttgaanaacaaagt	2373
Db	2472	cctcgcttttgaaagagcgacgtcgcctataacccgcgtcaagtaatgacvtgaagaacaaagta	2511
Oy	2374	tacacatactctactgtttaaagaagcctggggaaatgtagttctgccttgaacttcgg	2433
Db	2532	tactgcgtatctcctcaggttgaaagcttctcttgggggaataatgcttttaaatgatgttgg	2511
Oy	2434	tggagaagcctccgatttgccttagatagaagaatgctctatatttgagcgatcaatccctcat	2493
Db	2592	agctctctctcatctctatctctgaataacgtgcattgttttgtaacctatgctcccatat	2651
Oy	2494	gaatttcagttgtctatgcacatcagaagaagttttaagaacagcggaacagaagctcg	2553
Db	2652	caaatgtaatctgcacctataatcgtcaagaacagctctctcggaagaagytacagaaggaag	2711
Oy	2554	tgaatttggagaagtagccgctctgttgaaactcttgccctacacacggatccgaatttgataa	2613
Db	2712	atccttttgagaagcagcaacctcttcaattatcttgcctataaggggtgaagttagaga	2771
Oy	2614	ggaatcagactgcgcaagaatgcaagcagcaacatcgaactcttggtttatctcgtgtaactgt	2673
Db	2772	gtctctcgattgataagaaactctctctatgatacgcgacttatccctatgctcgtgactcat	2831
Oy	2674	tctgtagtaaccccgacgtgtacgaacaacactgcgaattagcgggtgatactcttgaaacct	2733
Db	2832	ccgcaatgatacccaatgacactacagcacttgtaaatcagcagcgactcttggaaacta	2891
Oy	2734	cggtagcaatttggcaagaagaacttagtctctcgtgcagggaaacattttgcttaa	2793
Db	2892	tgcgaataacttagcagcagaagcctcttgcaagtcgtgcaggaagtlactaagcctcttc	2951
Oy	2794	ctcaaatcttgaaagccttagcaaatcttcttttgattgacgttgggtc	2841
Db	2952	tctcatgttggaagtgctcgtgcgaagtttgccttggaagtctgtgatac	2999
RESULT	10		
ID	X06816		
XX	X06816	standard; DNA; 3200 BP.	
AC	X06816;		
XX	XX		
DT	26-APR-1999	(first entry)	
XX	XX		
DE	XX	Chlamydia pneumoniae surface exposed protein Omp4 DNA.	
XX	XX		
KW	XX	Omp4; outer membrane protein 4; surface exposed protein; antigen;	
XX	XX	infection; diagnosis; vaccine; atherosclerosis; asthma; ss.	
OS	XX	Chlamydia pneumoniae.	
XX	XX		
FH	XX		
FT	XX	Key Location/Qualifiers	
XX	XX	205..2991	
XX	XX	/*tag= a	
XX	XX		
XX	XX	W09858953-A2.	

PD 30-DEC-1998.

PF 19-JUN-1998; 98WO-DK00265.

PR 23-JUN-1997; 97DK-0000744.

PA (BIRK/) BIRKELUND S.

PA (CHRL/) CHRISTIANSEN G.

PI Birkelund S, Christiansen G, Knudsen K, Madsen A;

PI Mygind P;

DR WPI: 1999-105610/09.

DR P-PSDB; W88417.

XX

XX

XX Species-specific test for identifying mammals infected with

PT Chlamydia pneumoniae - comprises detecting antibodies specific for

PT outer membrane proteins of C. pneumoniae or nucleic acids encoding

PT these proteins

PS Claim 6; Page 35-40; 115pp; English.

CC This DNA sequence codes for the novel 98.9 kDa surface exposed

CC protein Omp4 (see W88417) of the human respiratory pathogen

CC Chlamydia pneumoniae. By generating antibodies against C.

CC pneumoniae outer membrane complex, a polyclonal antibody (PAB 150)

CC was obtained which reacted with outer membrane proteins. The

CC antibody was used to identify the genes (see X06816-27) encoding

CC Omp4-Omp1s proteins (see W88417-28) in an expression library of

CC C. pneumoniae DNA. The genes are situated in 2 gene clusters:

CC Omp12,11,10,5,4,13 and 14 in one cluster and Omp6,7,8,9 and 15 in

CC the other, and encode polypeptides of about 89.6-100.3 kDa and

CC about 56.1 kDa. The invention provides a new species specific test

CC for identifying mammals (including humans) infected with Chlamydia

CC pneumoniae. The test comprises detecting antibodies specific for

CC Omp4-Omp1s or detecting nucleic acid fragments encoding these outer

CC membrane proteins, especially by PCR. The proteins are also used

CC in the diagnosis of C. pneumoniae infection in mammals. The

CC nucleic acids and proteins can also be used in the immunization of

CC mammals, the nucleic acids being particularly useful as DNA

CC vaccines for effecting in vivo expression of antigens. The

CC vaccines may also prevent atherosclerosis and bronchial asthma,

CC which are possibly associated with C. pneumoniae.

SQ Sequence 3200 BP; 946 A; 708 C; 623 G; 923 T; 0 other;

Query Match 15.5%; Score 465.8; DB 20; Length 3200;  
Best Local Similarity 52.0%; Pred.No.5.8e-119;  
Matches 1372; Conservative 0; Mismatches 1177; Indels 90; Gaps 11.

OY 296 cctggacaacgcacagcgcacatccaaaaggcttttaacaacactaaaggcgattgact 355  
| | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 394 cctggaaaaagcacacccttcattctgacagtgttlttaagcaaacacgagcaatctaac 453  
| | | | | | | | | | | | | | | | | | | | | | | | | | |  
OY 336 ttacagaagtaaocgggaacctcttatgtttccaacacggtgagtcagggattagcagg 415  
| | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 454 tttctggggaacgcgatcagcttagcttaagtgcttataagatgctgcgaactcatgag 513  
| | | | | | | | | | | | | | | | | | | | | | | | | | |  
OY 416 gctgtgtttaacagcagaggtgtgtatataatcttaccaacggtttatagggtttcttcgcta 475  
| | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 514 gctgtgtcatctaa---cagcaataataagaatcttacctcttcgggttttctcttacy 570  
| | | | | | | | | | | | | | | | | | | | | | | | | | |  
OY 476 tccttattgcgtccctcctggaaagttcgaataactacacggcaagagcgcttagcgtct 535  
| | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 571 agtttgttctctctccctcagcaaacggttctactaagtcagggaagcct---tcctca 627  
| | | | | | | | | | | | | | | | | | | | | | | | | | |  
OY 536 accggtagcttgagtttgcacaaaaaatgctcagttgtgctcttcacgaaaaaactttcaag 595  
| | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 628 gcagggaagcgttaaatltagaanaatctcgtaaaccttgatgtcgtgggaatttttctact 687  
| | | | | | | | | | | | | | | | | | | | | | | | | | |  
OY 596 gataatggcgtgtgcatcacccgcaaaaactcttcaataaagggacgacaatgctcagct 655

Db 688 gcagatggtgagctatataaaggagctcttccctttaactgcacttcggagatgct 747  
QY 656 ctgtttttgaaataacctctcaaaagaagcgagccattcagacttcggatgccctt 715  
Db 748 ctttttagtaacaactctcatcaaaaaggagagcaattgctactacagacggcct 807  
QY 716 accattactggaacaacaggaggagctctctttctgcaatactcttcggatctgga 775  
Db 808 cgcataagcaataaacaagggtatgttagatccctatcatcaatagcgtctcagtlcaag 867  
QY 776 gctgcaatttttaagaagccctcgtagctatttcttaataagtcaaggttcccttatt 835  
Db 868 ggcctatcgatgtgaaagcgacgctcgtactatccgaacaacaattctctattttgaa 927  
QY 836 gacataaggtcacaagagcgagctctcacaacgvgggtatgtcaggaggtgctac 895  
Db 928 gggagatg-----cagcgaaacaactcgtgcgtagcgc 960  
QY 896 tgtgtcttaaaactagtaacagatcactaaggtcacccctcacttggaatacagatttactc 955  
Db 961 tgcacaaccaagcgagatgagctctccctgaacctgaataccttaacaataaagactctgac 1020  
QY 956 ttcagcaacaatacagcacacagcgaggagctatctatgtgaaaaaagctcgaactg 1015  
Db 1021 ttgtctcaaacgtagcagaacaacagcggtgcgccatccactgtcttaaaagctgacccct 1080  
QY 1016 gcttcggagagcttaccctcttcagtgaataatgttcaatgtaggttaacagctcctaaa 1075  
Db 1081 tctctcggaggcttcaacagagttctcaagaaataatgtctcatag---caactcctcaag 1137  
QY 1076 ggttgagccatagctatcgaagaatgtaggtggaattgagttatccgacgagatggtgac 1135  
Db 1138 ggggggtgtatcagcatcgatgcgtccagagagctcagctcttcttgacagagacgggaac 1197  
QY 1136 atgtctcttttaggggaatacagctactctaac-----tactcctggagcagatagaagt 1189  
Db 1198 attcacccttgtaagaaataaccccttacaacaacggaagtaaccgatactccctaaagctaat 1257  
QY 1190 agtagcacttaagaaacagagtgcaaaagatgacagcttggcttggctgctggagagcc 1249  
Db 1258 ggcgtacacatagaaatgaacggaaattccggaattacgggctgcttaaaaaacatata 1317  
QY 1250 atctactctatgatcccaataactacagatcatccacaacaggttaacagatgtcttaaa 1309  
Db 1318 atttctctatgctcccatcactcagaaggaacccca-----tcagcgctattigaag 1371  
QY 1310 gttaatgagctccggagagattctgcactcaatatagagggaacatcatccttcacagga 1369  
Db 1372 ataataacggctctcgtggagagcttcaatccatatacaagaacgattcttcttctgga 1431  
QY 1370 gaaaagttacagagacagagggccgcagattcttaaaaaatttacttgaaagctactacag 1429  
Db 1432 gaaacccctaacagagatgacttaaaagtgtcgacaattttaaatttcaatccaagcag 1491  
QY 1430 cctgttaacttcttcaggaagtactctatctttaaacaatgtagtgcgtgcagactcag 1489  
Db 1492 ccagttccctatccggaaggaagtattgtctcaaaaaggagctacatttagagagcag 1551  
QY 1490 gcaattcaacaacagagagttctcgtctgaataatgtagcagtaagaactcctagaacct 1549  
Db 1552 agcttccctcaaaaggcgggttctcctcgtcgagatgattccaagaaacgaattatataact 1611  
QY 1550 ---gctgatactagcacaacaacaattgtgcatatacactgcgtctatagacggtgca 1606  
Db 1612 acagcttgaggtatttaacatcacgaacctgagatcaatgttaaccttaagttcttaag 1671  
QY 1607 aagaagggcaaaaatagaaacaaagctagctcaaaaatctgacttaattctggaacacatc 1666  
Db 1672 cagcccgctcagccttaacagaaaagtgctccaataaagtgtcgtatcttggaagcctc 1731  
QY 1667 acttattggaaccgagcgagcttttatgaaatcatcagtttaagaactcctcagctc 1726  
Db 1732 aacctgatgatattgaagggaacattatgaagaatcatatgttctcagccatgacccagctc 1791  
QY 1727 taagacatcttaagctcacaagctctcgaactcgttaacagacccgag----- 1775  
Db 1792 ttcccttatataaaaaacaaggttgaatgctgtagatctatcagtgtgacatcagcagc 1851  
QY 1776 -tgactcagatccataaatgggtgagaaaatccatcagctacatcaggaactgtgggc 1834  
Db 1852 ctatccctgttccgtcgtgagatcccaattcagaatacaggtatcccaaggaacattgaaat 1911  
QY 1835 ccaattgttgggggagaagggttctcagac-----tgcaacctcaactgactaa 1888  
Db 1912 gttaatggaactcaggaatacagaataaataaagaagccacggcgaacttggacaaa 1971  
QY 1889 actggtcataatcccaatcccgagcgtatcgcgtcttgaatccctatagacttgaatt' 1948  
Db 1972 acagagattgttccagcccccgaagaataatcgcggtttagtaagcaatacctatgggga 2031  
QY 1949 gcaattagaatatagctctctccatlatcctatgtagagactgcaaaaggggttgcag 2008  
Db 2032 gtctttagtacaatcgcctctcgtcgaacagcttgtagagctgcggcaactggtatgaa 2091  
QY 2009 gggagccggtcttgggtggtggtgatactaaacttctccatagaatagataa 2068  
Db 2092 cacaacaaggttctcgggttccctccatgacgaacttccctgataagaactggaatgaa 2151  
QY 2069 acacgagcggggttcgccaatttgagtggcggtatgctacataagagaagaaacctataact 2128  
Db 2152 aatcgcaaaagcttccgctacactctgtgagtgactgcgtatcggtgaaagtgcactact 2211  
QY 2129 tgttcagataagatttcttgggtgctgctatttgcagctcttggagaagatagagactac 2188  
Db 2212 cctaagaagcaactatlaacttgccttgcctcgtcactcttgcgtagagaacaagattgt 2271  
QY 2189 ttgtagtaagaatcaaggtacagctcagcgagagacctcctcattacagcaacaacgaa 2248  
Db 2272 ttatgcgtacaacaacctcagaaactcagctcagtgacaggttcttctcaagcactcat 2331  
QY 2249 acctatatactcttctcgtgcaaaactacggccttgcgttgctatg----- 2297  
Db 2332 accctacaaccccaaacatttgaatagaagaagcaagatttctgatacagctata 2391  
QY 2298 -----ttccctacaagattcctgttctcttcttcaagaaaccttagccacccaatag 2350  
Db 2392 gaaaattccctcagggaaatcccccagcctcgtgagtccaaagttcgttcagcaattca 2451  
QY 2351 gataacgatctgaaaaacaaagatacaacatactcactggttaaaaggagcggtgggaat 2410  
Db 2452 gaaacacgtatggaacacgacatatactcattcgcagaatccgaaggttctctgggcaac 2511  
QY 2411 gatagttgcgtttagaattcgtggtggaagactccgatttgccttagatga---aagtgct 2467  
Db 2512 gaggtatagcgtgtgtgtgactcgcgcctagaccttcttgccttcccaaccacatcct 2571  
QY 2468 ctattgagcagtagctacgtcccttcagaatagcaggttgcgtatagacatagaagaagt 2527  
Db 2572 cttttcaagaaccttcaatccacagatgaagtcgaaatggttctatggttcaaaaaatgac 2631  
QY 2528 tttaaaagacagggaaacagaagctcgtgaatttggaaatgtagcgtctgtgtaattctgc 2587  
Db 2632 ttcttcgaagactcgtatgtagtgcgtgtgttttagtatggaagcgtcttaacctccg 2691  
QY 2588 ttaacatcgggatccogattgtatagaatacagctcgcgaagtgcagatgcaactta 2647  
Db 2692 attccgttggtgagaaatctcgtgcaagggtatccgagatcttcaacaactatgatctc 2751  
QY 2648 actcctgtatcacgtgtgactgttgcgtatgaaccccgagctgtagagaacaacatgga 2707  
Db 2752 taagatatttctgttcogagtctcatcgttaacaatcccaattcaacgagcactctgtg 2811  
QY 2708 attagcgtgattcttggaaaaacttcgtgtaacgaatttggcagaagaagctttagtctt 2767  
Db 2812 atgagccagactcttggaaaattcgcgttgacatcttccaagcaggaacttttactg 2871

OY 2768 cgtcaggaacccatttttgccttaacaatttgaagcccttagccaatttcctt 2827  
     ||| |||| |     || |||| |||| |||| |||| |||| |||| ||||  
 Db 2872 aggggtagcaacaactcgccttacaaactcgatcttgagcatltagcta 2931  
     ||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Oy 2838 gaattgcgttggtgatctgcgaattacaatgtagtagagagaataaccattca 2886  
     |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Db 2932 gaactcgcgtgagtcttcaagaactacaatagtagatgttgtagaccaactcgaattca 2990  
     |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

**RESULT 11**  
**ID** 261508 261508 standard; DNA; 3050 BP.  
**XX**  
**AC** 261508;  
**XX**  
**D7** 19-JUN-2000 (first entry)  
**XX**  
**DE** DNA encoding the CPN100394 polypeptide.  
**XX**  
**KW** CPN100394; Chlamydia infection; immune response; vaccine; ss.  
**XX**  
**OS** Chlamydia pneumoniae.  
**XX**  

<b>Key</b>	<b>Location/Qualifiers</b>
<b>FH CDS</b>	101..2938
<b>FT</b>	/*tag- a
<b>FT sig_peptide</b>	101..226
<b>FT</b>	/*tag- b

**XX.**  
**PN** WO200011183-A2.  
**PD** 02-MAR-2000.  
**XX**  
**PE** 18-AUG-1999; 99WO-IB01449.  
**XX**  

20-AUG-1998;	98US-0097187.
20-AUG-1998;	98US-0097188.
20-AUG-1998;	98US-0097189.
20-AUG-1998;	98US-0097190.
20-AUG-1998;	98US-0097191.
20-AUG-1998;	98US-0097196.
20-AUG-1998;	98US-0097197.
20-AUG-1998;	98US-0097198.
20-AUG-1998;	98US-0097199.

**XX**  
**PA** (CONN-) CONNAUGHT LAB LTD.  
**PI** Murdin AD, Oomen RP;  
**XX**  
**DR** WPI; 2000-224703/19.  
**DR** P-PSDB; Y69368.  
**XX**  

Novel antigens and corresponding DNA molecules that can be used to prevent, treat and diagnose disease caused by Chlamydia infection in mammals, especially humans -

Claim 1; Fig 13A-F; 201pp; English.

Z61502-09 encode Chlamydia pneumoniae polypeptides. The polypeptides are present in the bacterial membrane structure, in the external vicinity of the membrane structure, in the inclusion membrane structure, in the external vicinity of the inclusion membrane structure and in the cytoplasm of the infected cell. The polypeptides may be used to prevent, treat and detect the presence of Chlamydia infection and/or the presence of Chlamydia in a sample. The polypeptides may also be used to induce an immune response in a mammal. The vaccine vector comprising the polynucleotides is used to induce an immune response in a mammal. Antibodies directed against the polypeptides may also be used therapeutically to treat and/or prevent a Chlamydia infection.

Sequence 3050 BP; 897 A; 622 C; 695 G; 836 T; 0 other;

Query Match	13.0%	Score 389;	DB 21;	Length 3050;
Seq Local Similarity	50.7%	Pred. NO. 1.le-97;		
Matches 1422;	Conservative	0;	Mismatches 1270;	Indels 111;
				Gaps 16;
QY	164	gtctaccgagacggttttgyatcgaatgicgagtttcgaatgggaaataaataatgtaatttt	223	
DB	170	gtaacctccactacctaatactccctgaagatggtttatttgggggccaatacaataactttt	229	
QY	224	tcaagttcttgagatgcaggaagaatgcttgyaactactctctcatttaaggyaatgtcaat	283	
DB	230	tctccgaaatctacacaggaatgctgcgaagaactactctctctccacagggagagt---	286	
QY	284	ctagaanaatatcttcgggacacgacagcaatccacaanaagctttttaacaacacag	343	
DB	287	cgtatatatagatccgggaaaggtggtgttcatttaccagaatctgtcttgtgaacgtct	346	
QY	344	ggcgatttgactttcacaaggtaacggggaactctctatgttccaaacggtgagatgcagg	403	
DB	347	ggcgattcttaccttttttagtcatbgtaaataccctaaagttcttcgtgtagatgcaggt	406	
QY	404	actgcagagagggcgctgcgtttaacagcgcgctgtgagtaatatccacagtttatggg	463	
DB	407	gtcaatatccgggtgtgccat-----gtacaaggaaagtaagaatttaagcttccacgat	460	
QY	464	ttttctcgctactcttttattgcyctctcctctggaagtttcgaatactccggcaaggagcc	523	
DB	461	ttctctctctcgtgtatcacagaatctcccaaatccgctgttactacaggaagaagtacg	520	
QY	524	gtttagctgctctacggggtgaagcttggtttggcaaaaatgtccagttgtcttcaagcaa	583	
DB	521	ctagtcagatct--aggtgcagctccaaactgtgaataataaacaactctgttctttacaagc	577	
QY	584	aactcttcaacaggaataatgycgggtgcataccgcaaaacactcttcaataacagggact	643	
DB	578	aatgcctctgtccgaagatggtggtgtgtttaaggaaactctgttattcagggaatc	637	
QY	644	acaaatgcagctcgttcttctgnaaataccctctccaaagaagaacgggaccattcagact	703	
DB	638	aaaaatagtcgatttttggacaataatcatcttcgnaaaaagagggcgatctccacg	697	
QY	704	tccggtgcccttaccattactcgtgaaaccaagggggaagctctttttctgcacaatacttc	763	
DB	698	actcaaggaacttaaccatagagaataacttaaggagcgcgaagttccaatgaaaaacaagca	757	
QY	764	tccggaattcgtgagctgcacatttttacaagaagcctcgtgtaatttttcaataatgcataa	823	
DB	758	gtgacctcagagggcgctttagattagagacggtcttaattcatcactcgtgcacacatgag	817	
QY	824	gtttccttatttgcacaataaggtccagga---gcgagctccccaacaacgggggatatg	880	
DB	818	tgcataatttccacaanaataaagactcttggaatgctgcgaatggcgggaccataaatgvc	877	
QY	881	tcaaggaaggtgcatactgtgcttaaaaact-----	910	
DB	878	tcaaggagaccttaacttactgataatacaactcttctgttacttcaagaanaatagcacatg	937	
QY	911	-----agtaacagatactaaggttcaacctcactctggaataatcaagatgtta	952	
DB	938	cagagatgtgtgagcttctgttagcacaggaaccataagcattaccggtgaggtattcatc	997	
QY	953	cccttcagcaacaataactatgcacaacgcgggagaggtactatctatgynaanaagctcgaa	1012	
DB	998	aatgtcataganaaataacttcaganaaaaagagagaggtatttgcagcttctctcaag	1057	
QY	1013	ctggtctccggaggaacttaccctatcagttagaaatagttgcaatgagtaagcagccct	1072	
DB	1058	atttggagagggcagggggtgcgtctcttcttataaacytaagatcatatgcacccct	1117	
QY	1073	aaaggtgagccatactgcatcgaagatagtgtgggaattgagttatccgcgcgatagtgt	1132	
DB	1118	ctaggaggtgtcatattttatacaacaaaggaagactcttgcagcttcttcaatcacaaggagq	1177	



QY 1133 gacattgtcttttagggatacaagtcacttctactc-----ctggagcgataga 1186  
 DB 1178 gatattcggtttcggagggaatcagatcaacacagctccaatgctacccaataagaga 1237  
 QY 1187 agtaagtacgttaggaacgagtgcaaaagatgacagcttgcgttcgtcgtgtaga 1246  
 DB 1238 aatgtaattcacctcgaagcgacccggaagttagcggagctgcgcgaagcaagtaac 1297  
 QY 1247 gccatctctttatgatccccaatacagatcatccaacaacagttacagatgtcta 1306  
 DB 1288 gctactctcttctatgatacccatccac---caacgatacgggagcaagataactta 1354  
 QY 1307 aaggttaatgagactccgcagatctcgaatataataagggaaatattcttcca 1366  
 DB 1355 cgtatcaatgaggtcagcgcaaat-----caaaagctctcggatcttatgtatttct 1408  
 QY 1367 ggaagaagaattatcagaagaacagagccgagattctaaanaattactactcgaaacta 1426  
 DB 1409 ggaagagagattgtc---gacagcgaaagctataagctgaanaattactactcgagataac 1465  
 QY 1427 cagcgtgaactctcttcagggagtaactatcttaaacatgagtgagctctgcagact 1486  
 DB 1466 cagcgtgcactctttagtgagggagagcttagtacttaacagggagtgagcttgcataca 1525  
 QY 1487 caggtcatcactcaacagcgagatctcgtctcgaatgagcgtaaggaactactctaga 1546  
 DB 1526 caagatattctcgagagccagaatccagcttcttggatcgctggagactcttaaa 1585  
 QY 1547 ccttct---gacttagcacatataaactatgctacttaacatcagttctatagacgt 1603  
 DB 1586 gctctacagagaatattgtcatcaacaattatcataatgcccgaataccaattacgga 1645  
 QY 1604 gcaagaagggaaanaatagaaaccaaagctacgtcaaaaatcgcgaattattatgagac 1663  
 DB 1646 aagatctcatcaatattgttagcttcagcgcgagaataagaacattaccctcaacaggaac 1705  
 QY 1664 atcacttatatggaccgagcgagcagcttctatgaaatacatagtttaagaaatccctag 1723  
 DB 1706 ttgactctgtaaatgagatgtagtcttctatgaaacacatccttgcaagctctcaca 1765  
 QY 1724 tctacgacacatcttagagcccaa-----agcttcggaactgtaaaacagacc 1771  
 DB 1766 gactatagctttgtaaggtatattcccgagagcggaagagactaattactcaagatgt 1825  
 QY 1772 gcaagtgcctccagatccctataatggtgtagaaatcccatcagcgatcaggaactgtg 1831  
 DB 1826 tctcagaagcctctgaaagtagctccttctagaccacattatgctataagaaacttgg 1885  
 QY 1832 --ggccaatgtttggggagaggggcttctagact---gcaacctcaactgact 1885  
 DB 1886 aatgtgcaagtcattccaggaacggaactcaacgagccagccgaactatagaaatgtgt 1945  
 QY 1886 aaaaactggtatattcccaatcccgagcglatagctctttagtcccatagctatgtg 1945  
 DB 1946 cggacagagataccttcgaatcccgaaagcgcaagatctttagtcccatatgctgtg 2005  
 QY 1946 aatgcatatagatatagctctctcccatattctatgaggaactgcgaacgaaggtgtg 2005  
 DB 2006 ggttctttgtgcatcagcggtgctatccaagaatcatgtaaatagtagccaactctta 2065  
 QY 2006 cagggaagacggttctttgtgtgtgattatccaatttccatagatagatgtaca 2065  
 DB 2066 tgcagagaaacgggaggtcggggagctgtaattgtcattcccatatagaga---taaa 2122  
 QY 2066 aaaaacgcgagcggttccgcaattgagtgcggtatagcatagaggaacactaat 2125  
 DB 2123 ataatagacacggcatcgcatagcggtgtcgttatcttggtggagttggacctcat 2182  
 QY 2126 actgttcagataagattcttgatgctgattttgcaagctcttggagagatagagac 2185  
 DB 2183 gcttttctcgaatcagataatgcgagcttcttgcagccttctcagtagagataaagac 2242  
 QY 2186 taactttagtaagaatcaagatcaggtctacaggaagaaactctctattacagaaacac 2245

DB 2243 taagtagtaccacaanaactctggaactagctacaggggtcgtatttcttggagatcc 2302  
 QY 2246 gaaacctatctctctctcttcccttgcgaactacagcgcttgcgttgcgttctatgtctaca 2305  
 DB 2303 ctgaggtttagaagtcacaggg---attctatcgtatagctccccaagaagctgtct 2359  
 QY 2306 gagatctcttctcttctttaggaacacttagctaaaccatacagatacagatcgtaaa 2365  
 DB 2360 aaccaagtcgtcactatagatagtcagttgtcttacaagccatagaataatgatagaa 2419  
 QY 2366 accaagtatacaacatactactgtttaaggaagcgttggggaggaatagatgttcgtta 2425  
 DB 2420 accaataacacgacatataccagaagctcagggatccttggcaaatgagtgttggctt 2479  
 QY 2426 gaattgggtggagaagatcccgatttgcgttaagtgaagtgcctctatgagcagatag 2485  
 DB 2480 gagttcgagcgactacataactactaccctaacagtaacttcttatttgattactactct 2539  
 QY 2486 cccttaatgaattgaggttgcgttctatgacatacgaaggttttaagaacagagaa 2545  
 DB 2540 ccgttctcagcgctgagtgacacatagctgtcccggaagactcaagaagacagagagt 2599  
 QY 2546 gaagctcgtgaattggagaatgagccgcttcttgaactctgcttaccatcggaatcga 2605  
 DB 2600 gaggttcgtcacttactagcggagatccttcaatttagcagttccatgtcggtgag 2659  
 QY 2606 ttgataagaagtaacgagctcgcgaagatgcaacgttcaatccactcctgtgtatctgtg 2665  
 DB 2660 ttggagagatttctcagactgttaaaaggagatctttagaacttacccttgcgttagtct 2719  
 QY 2666 gatctgttgcgttagtaaacccgactgttacgacaacacgcaatgaggtggtatctgg 2725  
 DB 2720 gatgtgattcgcgaagaagatcccaagagcagcgcaacatg---gctagtggagctcaggt 2776  
 QY 2726 aaaaactcgcgtgagcaatttggcaagacagcttagtccctcgtgcgaaggaacattt 2785  
 DB 2777 agcacccaaggaanaactctctccagacaagatatacactgcgtttaggaacactgt 2836  
 QY 2786 tgccttaactcaaattttgaagcctttagccaatttcttctttagatggcggtgactc 2845  
 DB 2837 ctcatataatcccgatattgaggtgttcagtcagtcaggaactatgagtcggggtcct 2896  
 QY 2846 cgcatttcaatgtgacttaggagcaaatatccaattctaat 2888  
 DB 2897 cgtaatatacatcaatctcgggggttaaatatccgattttaat 2939  
 RESULT 12  
 X06827  
 ID X06827 standard; DNA; 2838 BP.  
 XX  
 AC X06827;  
 XX  
 DT 26-APR-1999 (first entry)  
 XX  
 DE Chlamydia pneumoniae surface exposed protein Omp15 DNA.  
 XX  
 KW Omp15; outer membrane protein 15; surface exposed protein; antigen;  
 KW infection; diagnosis; vaccine; atherosclerosis; asthma; ss.  
 OS Chlamydia pneumoniae.  
 OS  
 PN W09856953-A2.  
 PD 30-DEC-1998.  
 PD  
 PF 19-JUN-1998; 98WO-DK00266.  
 PF  
 PR 23-JUN-1997; 97DK-0000744.  
 PR  
 PA (BIRK/) BIRKELUND S.  
 PA (CHRI/) CHRISTIANSEN G.

XX Birkelund S, Christiansen G, Knudsen K, Madsen A;  
PI Mygnd P;  
XX WPI: 1999-105610/09.  
DR P-PSDB: W88428.  
XX  
XX Species-specific test for identifying mammals infected with  
PI Chlamydia pneumoniae - comprises detecting antibodies specific for  
PI outer membrane proteins of C. pneumoniae or nucleic acids encoding  
PI these proteins  
XX  
XX Claim 6, page 70-71, 115pp: English.  
XX  
XX This DNA sequence codes for the novel surface exposed protein Omp15  
CC (see W88428) of the human respiratory pathogen Chlamydia pneumoniae.  
CC By generating antibodies against C. pneumoniae outer membrane  
CC complex, a polyclonal antibody (PAB 150) was obtained which reacted  
CC with outer membrane proteins. The antibody was used to identify  
CC the genes (see X06816-27) encoding Omp4-Omp15 proteins (see  
CC W88417-28) in an expression library of Chlamydia pneumoniae DNA.  
CC The genes are situated in 2 gene clusters: Omp12, 11, 10, 5, 4, 13 and  
CC 14 in one cluster and Omp6, 7, 8, 9 and 15 in the other, and encode  
CC polypeptides of about 89.6-100.3 kDa and about 56.1 kDa. The  
CC invention provides a new species specific test for identifying  
CC mammals (including humans) infected with Chlamydia pneumoniae. The  
CC test comprises detecting antibodies specific for Omp4-Omp15 or  
CC detecting nucleic acid fragments encoding these outer membrane  
CC proteins, especially by PCR. The proteins are also used in the  
CC diagnosis of C. pneumoniae infection in mammals. The nucleic acids  
CC and proteins can also be used in the immunization of mammals, the  
CC nucleic acids being particularly useful as DNA vaccines for  
CC effecting in vivo expression of antigens. The vaccines may also  
CC prevent atherosclerosis and bronchial asthma, which are possibly  
CC associated with C. pneumoniae.  
XX  
XX Sequence 2838 BP: 837 A: 591 C: 648 G: 762 T: 0 other:

Query Match 12.6%; Score 378.4; DB 20; Length 2838;  
Best Local Similarity 50.5%; Pred. No. 8.8e-95;  
Matches 1415; Conservative 0; Mismatches 1276; Indels 111; Gaps 16;

QY 164 gctaccggagacggttttgatccaagtgcgagttcgcgtgggaataaattgtaattt 223  
DB 70 gaaccactccactaatctcgaagtgtgtatttgggaggaacaatactattt 129  
QY 224 tcagttcgtgagtcaggaagatctgaactactactctttaaaggaattgtaatt 283  
DB 130 tctccgaatctcaacaaggatgctgcgaagaactactctcttcacaggaaggt---t 186  
QY 284 ctagaanaatctctggaacagcagcaacaatcacaagaagctgttttaacaacctaa 343  
DB 187 ctgttatagatcgagggaaggtgttcaattacagaactgtgtgtgaactgct 246  
QY 344 ggcggtttgactttacaggtgaaggaactctcattgttccaaaggtggaagcagg 403  
DB 247 ggcgactctacattttagtaattgaataaccctaagttctcgttagatgcagt 306  
QY 404 actgtagaggggctgcgtttaacagcagcgtgtgtagataatctcaacgttttaagg 463  
DB 307 gctaatatcggtgtgcatc-----gtacaaggaagtaaaatttaagcttccagat 360  
QY 464 tttcttcgctatctttattgtctctcctcgtgaagttcgataactaccggcaagaagcc 523  
DB 361 ttccttctctgtgtatcacagaatctcaaatcgcgtgtagtcacaggaagtagtc 420  
QY 524 gtagctgctcaaggagctgtgttgacaaaatgttaagtttctcttcgcaaa 583  
DB 421 ctatgcagttca---gggacagtcacaactgcgaagataaacaactgttcttcaagc 477  
QY 584 aactttcaacgataatggcgtgtatcaccgcaaaaactcttcaattcaacaggaact 643

DB 478 aatgcctctgcgaagatggtggtgtatgaaggaactcctcttgattcagggaatc 537  
QY 644 acaatgtaagctgttttcttgaataatccctccaaagaagcggaaccattcagact 703  
DB 538 aaaaatagtgatttttggacaataatcattcttcgaanaaaagggggtgctccagc 597  
QY 704 tccgagctccctaccattctcgaacccaaggaagctctcttcttccgaatacttct 763  
DB 598 actcaagcaccaccatagagaataactttagggcgtcaaatgttcaatgaataaagca 657  
QY 764 tcgattctgtagctgcaatttttacagaagcctcgtgtactattcctaataatgctaa 823  
DB 658 gtgacctcaaggagcgcttagatttagagcgcgctcatcattcactcgcaaccatgag 717  
QY 824 gttcctctattgcataaagaagtcacagga---ggagctcctcaacaacgggggata 880  
DB 718 ttgatatcttcacaaaataaagactcttggaaatgctgaaatgycggacataatgct 777  
QY 881 tcaggaagtgctatctgtgtctataaact----- 910  
DB 778 tcaggcgagcccaaatattacgataacactcttgttactcaagaataagcacaatg 837  
QY 911 -----agtaacagatactaaaggtcaccctcactggaaatcagaatgta 952  
DB 838 caggatgtgtgagcttgtgtgtagacagagcaaccataagcattacggttagttctatc 897  
QY 953 ctcttcagcaacatactatcgaacacagcggaagagctactcttggaaaatcgtgaa 1012  
DB 888 aatgtgaaggaataatctcgcagcaaaaaggaagacgatttctgaagcttctcgaag 957  
QY 1013 ctggtctccgaggaacttaccctatcagtagaatagtgtcaatgtagagtagacgtcct 1072  
DB 958 atttggagggcgagggagcgccctcttcttaataacgafgactcaatgcccacct 1017  
QY 1073 aaaggtgagccacaagctatcgaagaatgtaggggaattgagttatccgcgcaatggt 1132  
DB 1018 ctaggaggtgtccattttatcaacacaggaagatcccttcagcttccactcaaggagg 1077  
QY 1133 gacattgtccttttagggaatacagtcactctcactactc-----ctggagcaagata 1186  
DB 1078 gatattcgattcgggggaatcagagttcactctcaacacagctcaatgttaccacaaag 1137  
QY 1187 agtgcgtacgacttaggaacgagtgcaagaatgcaagcttgcgtctgctgtgtgata 1246  
DB 1138 aatgttaatcaccctcgagagcaccgcgaatgtagcgagcttgcgaagcaaggtaac 1197  
QY 1247 gcatctactcttctgtatcccatatcactacagatcaccacaagaattacagatgctta 1306  
DB 1198 gctatctattctctatgatacccatlacac---caacgatacgggaggaacgataactta 1254  
QY 1307 aaagttaatgagactccgagatctcgcactcgaataatatacaggaacatcatctaca 1366  
DB 1255 cgtatcaatggggtcagtgacaat-----caaaagctctcggatctatagttttct 1308  
QY 1367 ggagaaaagtatacagaacagagcgagatctaaanaattcttactcgaagctacta 1426  
DB 1309 ggagagagattgtc---gacagcagaagctatagctgaanaattcttactcgaagtaac 1365  
QY 1427 cagcctgttaactcttcaagaggtactctatctttaaatactgagagctctgcgaact 1486  
DB 1366 cagcctgtcaacttagtagagggaggtctgaacttaacaaggaaggtgacttgcataca 1425  
QY 1487 cagcctcaactcaacagcgagatctcgtctcgaataatgtagaggaactcaactagaa 1546  
DB 1426 caagattctcgcagagcgcaagatccagctctctttagattggagactcaattcaaa 1485  
QY 1547 cctgct---gatactagcaccataaacaatttggctatttaacatcagttctatagcgt 1603  
DB 1486 gcttctacagaagatcgtatcacaatcattcatctttaaattgcccgaataccattagga 1545  
QY 1604 gcaagaagagcaaaaatagaaacccaagctcagtcgcaaaaatctggaattatctggagac 1663  
DB 1546 aagatccaatcaatattttagctttagcagcgagataagaacattacccataacaggaac 1605

```

QY 1664 atcaattatigaccgcagcgacggtttcttgaaatcatagtttaagaatccctag 1723
Db 1606 ttgaagactgttaaaagcagatgagcttctgtatgagaacacacttgcagacctc 1665
QY 1724 tctcagacatctttagagctca-----agcttcgagacgttaacagaccc 1771
Db 1666 gattatagcttgttaaaagtatctccagagagcgaggagacataactaactaa 1725
QY 1772 gcaatgaactccagatccataaattggtagaatccatccatcgatcaggaactt 1831
Db 1726 tctcagagactctcttgaaagtagctctctctcagacacattatgctatcaga 1765
QY 1832 --ggcccaattgttgggggagaggggcttctagact---gcaaccttaactg 1885
Db 1786 aatgtgcagatccatccaggaacgggaactcaacccagacgaatttaga 1845
QY 1886 aaactgcatatctcctcaatcccgagcgtatcgctctttagtcccttaagctta 1945
Db 1846 cggacagatccctcctcgaatcccgagcaggaagattttagtcccaatagcctt 1905
QY 1946 aatgattatagatattagctctcctcattatcttatgagactgcaacgaa 2005
Db 1906 ggttcttctgtatcagcagcgtctatccaaagaatcatgtaaatagagcca 1965
QY 2006 caggagagacgctgtcttgggtgtgtgtgattatctatctctccaaaga 2065
Db 1966 tgcagagacggggagctgtggggagctggaattcctatcccaataga 2022
QY 2066 aaacacgacgcgggttcgacatttgagtgcggtatgtcatagaggaa 2125
Db 2023 attaatgagcagcgtctcagcagtagcgtgtgtcttctgtgggagtgga 2082
QY 2126 actgttcaagataagattctttagtgcgcatcttgcagcttcttgaa 2185
Db 2083 gcttctctgtgtcgtatagataaatgcggtcttctgcagcctcctcag 2142
QY 2186 taactttagtgaagaaatcaaggtacaggtcaggaactcctatccacacaa 2245
Db 2143 tagttagatccaataatcatagtaagctacagcaggggtcgtatctt 2202
QY 2246 gaacacatattctctctctctcgtcaaacactagcgtctgtctgtct 2305
Db 2203 ctgaagtttgaagctcaccaggg---attctatctatagctcctcaga 2259
QY 2306 gagaattcgttctctctcttcaagaaaccttagctacacacatcggt 2365
Db 2260 aaccaagctgctactatagatagatgagtttcttcaagcagcatagaa 2319
QY 2366 accaagatatacaatactcctactgttaaggaagctggggagatgata 2425
Db 2320 accaataacacagacatccacagagcaggtatccttggcgaatgtgt 2379
QY 2426 gaattcggtggaagagctcgagatttcttgatgaagaagtctctattt 2485
Db 2380 gattttggagcgactacatactctacacctaaagacttttatttgg 2439
QY 2486 ccctcatagaattgcaagtttgtctatgcatcaggaaggttttaagacag 2545
Db 2440 ccgttctcagcgtcagtgacactatgctcaccaggaactccaagaaga 2499
QY 2546 gaagcgcgtgaatttggagtaacgtctctgtgaatccttgacctac 2605
Db 2500 gaagttcgttaacttctcagcgagatcttcaatttagcaatttgc 2559
QY 2606 ttgtataagaatcagatccagagtgcaagcttaacaaactaacttgg 2665
Db 2560 ttgtgagagatttccagactgttaaaaggatattatgaacttaacct 2619
QY 2666 gatctgttcgtatgaaccgcagctgtacagacacactgcaattagcg 2725
Db 2620 gatgtatccgacaagatcccaagagcagcgacaacttg---gctagt 2676

```

```

QY 2726 aaaccttcggtacgaatttggcaagagaagctttagtctcttctgca 2785
Db 2677 agcccccagggaaacaactctctccagacagatlaaactgtttagga 2736
QY 2786 tgccttaactcaaattttgaagcctttagccaatttcttcttgaattgc 2845
Db 2737 ctcaataactcgtgaattgggtgtgtcagcagggacttatggaatt 2796
QY 2846 cgcattacaatgttagcttagagagcaaataccattctaa 2887
Db 2797 cgtaatataacatcaatctcgggggttaataaccgattttaa 2838

```

```

RESULT 13
X06819
ID X06819 standard; DNA; 2526 BP.
XX
AC X06819;
XX
DT 26-APR-1999 (first entry)
XX
DE Chlamydia pneumoniae surface exposed protein Omp7 DNA.
XX
KW Omp7; outer membrane protein 7; surface exposed protein; antigen;
XX infection; diagnosis; vaccine; atherosclerosis; asthma; ss.
XX Chlamydia pneumoniae.
XX
PN MO9858953-A2.
XX
PD 30-DEC-1998.
XX
PE 19-JUN-1998; 98WO-DK00266.
XX
PR 23-JUN-1997; 97DK-0000744.
XX
PA (BIRK/) BIRKELUND S.
PA (CHR/) CHRISTIANSEN G.
PI Birkelund S, Christiansen G, Knudsen K, Madsen A;
PI Mygdin P;
PI
DR WPI: 1999-105610/09.
DR P-PSDB; W68420.
XX
PT Species-specific test for identifying mammals infected with
PT Chlamydia pneumoniae - comprises detecting antibodies specific for
PT outer membrane proteins of C. pneumoniae or nucleic acids encoding
PT these proteins
XX
PS Claim 6: Page 49-50; 115pp; English.
XX
CC This DNA sequence codes for the novel 89.7 kDa surface exposed
CC protein Omp7 (see W68420) of the human respiratory pathogen
CC Chlamydia pneumoniae. By generating antibodies against C.
CC pneumoniae outer membrane complex, a polyclonal antibody (PAB 150)
CC was obtained which reacted with outer membrane proteins. The
CC antibody was used to identify the genes (see X06816-27) encoding
CC Omp4-Omp15 proteins (see W88417-28) in an expression library of
CC C. pneumoniae. The genes are situated in 2 gene clusters:
CC Omp12, 11, 10, 5, 4, 13 and 14 in one cluster and Omp6, 7, 8, 9 and 15 in
CC the other, and encode polypeptides of about 89.6-100.3 kDa and 15 in
CC about 56.1 kDa. The invention provides a new species specific test
CC for identifying mammals (including humans) infected with Chlamydia
CC pneumoniae. The test comprises detecting antibodies specific for
CC Omp4-Omp15 or detecting nucleic acid fragments encoding these outer
CC membrane proteins, especially by PCR. The proteins are also used
CC in the diagnosis of C. pneumoniae infection in mammals. The
CC nucleic acids and proteins can also be used in the immunization of
CC mammals, the nucleic acids being particularly useful as DNA
CC vaccines for effecting in vivo expression of antigens. The
CC vaccines may also prevent atherosclerosis and bronchial asthma,
CC which are possibly associated with C. pneumoniae.

```

XX Sequence 2526 BP; 696 A; 495 C; 597 G; 738 T; 0 other;

Query Match 11.3%; Score 340.2; DB 20; length 2526;  
Best Local Similarity 51.18; Pred. No. 3.2e-84;  
Matches 1015; Conservative 0; Mismatches 913; Indels 57; Gaps 7;

QY 941 aatcagatgttactcttcagcaacaatatacatgcagcgaggagagatcatctgtg 1000  
DB 559 aataagctcccttcttcttatttgaaatagttcttcaacgcyggcgagatctacc 618  
QY 1001 aaaaagctcgaactggtctccggaagacttaccctatcagtagaagatggtcattgga 1060  
DB 619 aaaaactcaacatctctcgttggtgggaacctatttccagggagatacagcgctaacg 678  
QY 1061 gttacagctccctaaaggtgagagcatagctatcgaaagatagtggaattgattatcc 1120  
DB 679 g---ccgcggttaaaagaggtgcctacgcgacttcgcaacctatccatttct 735  
QY 1121 gccagatgtgtgacatctgtctttttagggaatacagctactctactcctgagag 1180  
DB 736 ggagacaggtgcgacatttctttgaaagcaatagataggtacaggaacgcgtct 795  
QY 1181 aatagagtagtactgcacttaggaacaggtgcaaaatgacagcttgcgtctgct 1240  
DB 796 cat---agtgctattgatttaggaactagcgttaagataactgcttactgctgcgcaa 852  
QY 1241 ggtagagcatctactctatctatgcttcccaactacagagcttcccaacagttacagat 1300  
DB 853 ggacatacagatatactacttattatgataccgactactgtaacagagatcgactgtgctg 912  
QY 1301 gcttaaaagttaatgagactcccgagacttctgcatacaatatacaggaacatcatc 1360  
DB 913 gcttcaatattaatagccctgactagtagagatacaagagatagtcggaaccatagtc 972  
QY 1361 ttcaacagggaaagtatcagagacagagcgccagatcttcaaaaacttctactcgaag 1420  
DB 973 tcttcgagagaaagctcgcgagagagcaagactaaagatgagaagaccgaccttataa 1032  
QY 1421 ctactacagccgttactcttcaagagtagtactcttctttaaagaatgagtagtactg 1480  
DB 1033 ttaactcaaaagtgtctttaaataatgagactgtgatttaaaggtatgctgtttaa 1092  
QY 1481 cagactcaggaacttcaatcaacagagcagattctcgtctcgaataatgagcagtaggaactac- 1539  
DB 1093 agtgcgaacggttctctcaggaatgcgaactcgaattgattatgattaggaagctg 1152  
QY 1540 -tctagaaacctgctgatactagcacataaacaattgtcattacaactcgttctata 1597  
DB 1153 ttggttcgcaaacacgcgaaggtatcgagttacgaatttggaaattatataagactcttc 1212  
QY 1598 gacggtgcgaagaagcgaagaataatagaaacaaagctacgltcaaaaactgtgcttact 1657  
DB 1213 aggaagagggaaaaaagataaactcagtgctgcacagcgtccagaagaatttctgataagt 1272  
QY 1658 ggaacataacttatttgaccccgagcgacggtttatagaaatcatagatttaagaat 1717  
DB 1273 cgtccgttgcgtactgcgaattagcgatgagaggttttaccaaaatggtcttggatgg 1332  
QY 1718 cctcagttcctagacactcttagagctcaagagcttctgagacgttaacagcacccgagtg 1777  
DB 1333 gacacattcctatgagtggaattcttgagttagatgcggtggaagaaactcgttatttctgca 1392  
QY 1778 actccagatcctataatggtgtagaatacttcaatcagctatcaggtgaactggggccca 1837  
DB 1393 gattctcgcagataaagtctgtacaaatctccgtatgctatcaggaagggagcaaac 1452  
QY 1838 attgtttggggagcagggctctcagactgcaacctcaactggtctaaactggctat 1897  
DB 1453 aatt-----ggtctactatgataagaagctacttcttctggtgcaagaagaatttt 1506  
QY 1898 attcctaaccgagaggtatcgctctttagtcccaatagctttaggaatgcatattata 1957

DB 1507 aatccacactgctcagcaggaaggctccggttagtctcctacttcttcttgggtctttata 1566  
QY 1958 gatattagctctctccatattcttattgagactgcgaacgaagggtctcagggagacgt 2017  
DB 1567 gatgttcgctccctccaaaatttataagctagtagtactgaagggtgcctcctacgaaga 1626  
QY 2018 gtttttggtggtgggttatactactcttcccttccctaaagatagatgaacaaaacagagc 2077  
DB 1627 agattttggttcgagcgacttcccaatttcttgcataaggacggtcgtgaaatacaagg 1686  
QY 2078 ggtttccgcaattgagtggtcggttattgcatagaggaagaaacctacatactgttcaagat 2137  
DB 1687 aaattccgcatagtgagtgaggtgctgtagaggtgcataagcagagatgcgggtggt 1746  
QY 2138 aagattcttagtgctgcatttctgtagctctttgaaagagatagagacttctttagct 2197  
DB 1747 gataccctgtcctcgtggttttgcagctcttgcggttgcaaaagactactttagaat 1806  
QY 2198 aagaatcaaggtacagctcagcaggaactcctatccagcaacagaaacctatac 2257  
DB 1807 accaatctcgaacaaagactcagcagatctttagcttgcagcaagatgcttccctatac 1866  
QY 2258 tctcttcccttgca-----aactagcgcttgttcgttcttagt 2299  
DB 1867 tctgtgtgagatctcctttagagagagagacccgcgaagatcctgttgccttagt 1926  
QY 2300 ccttagagatctcgtctctcttcaagaaactctatccatccacatagagataagat 2359  
DB 1927 tccaaagactctgcgtctcttcttcaagcagctttagcagctcagcaacagatcatgc 1986  
QY 2360 ctgaaaccaaagtatacaacatatactctgtttaaaggaaactggtgggagatagatttc 2419  
DB 1987 atgaagacggaggtctctaccccccccccccccccgacgctcgcagcagatcatactctgg 2046  
QY 2420 gctttagaatctcgtgtgagagactccgatttgcctatgta-----a 2461  
DB 2047 ggaagataatgtctcgtgagagcttggaactcagcaggttgcgttgaataacacgcggc 2106  
QY 2462 agtgcctatttgagcagtagctcctcctcaatgaaatgagctgttgcattatgacatag 2521  
DB 2107 agagatttctccgagatgatacctccatcttgaagaagccagcgttactcgcggcaa 2166  
QY 2522 gaaggttttaaagaaacagggaacagaaactcgtgaaatttggaaagtagcgtcttgat 2581  
DB 2167 gatagcttgttgaactagagagctacagctcgtgtgttttagatctgcacotctataac 2226  
QY 2582 ctgcttaactatcggatccgacttctgataagaaatcagaactgcgaagatgcaacgttac 2641  
DB 2227 ctgagatctcctctgaaatcaagttagaagaacggtttgcagagcaatat-----tat 2280  
QY 2642 aatcctaactctgttatacgttgaactgttctgttgaatacccgagcttgaagcaaca 2701  
DB 2281 catgttgaagcagatbattctccagatgttgcgtgaagaaacccaatgtagcgtacc 2340  
QY 2702 ctgcaaatlagcgtgacttcttgaaacacttcggtgaagaaatttggcaagaacagctta 2761  
DB 2341 ctacttccaacaaaggaggttgaagacaaaggcttgcgaacttgaagcagagcgttgt 2400  
QY 2762 gtccttctgacaggaacacatttgccttaactcaaatatttgaagccttagcaatt 2821  
DB 2401 atgttctggagcctcaggttttcgaactcttggaggtgcagcagagacttctggagactt 2460  
QY 2822 tcttttgaatgagtggttgcactcgcgaattacaaatgtagaacttggagcaaaataacca 2881  
DB 2461 ggccttgaatggcggtgtcttctcgtatgataatgtagatgcggtgagcaaaatacaa 2520  
QY 2882 ttctta 2886  
DB 2521 ttta 2525

RESULT 14





CC protein gene. Infection by Chlamydia can result in respiratory tract  
CC diseases such as bronchitis, sinusitis and pneumonia, asthma and  
CC atherosclerosis. The gene, protein and antibodies can be used as  
CC immunogens to induce an immune reaction in humans which has the effect  
CC of vaccinating the person. They can also be used to diagnose and treat  
CC those infected with the parasite.

XX Sequence 3150 BP; 886 A; 760 C; 622 G; 882 T; 0 other;

Query Match 8.1%; Score 243.4; DB 21; Length 3150;  
Best Local Similarity 49.7%; Pred. No. 2,3e-57;  
Matches 1002; Conservative 0; Mismatches 911; Indels 102; Gaps 11;

```
QY 914 acagatactaaagtcacccctcctcgtgaatcagatgttactcttcagcaaatacatcg 973
Db 1079 acaggtgttgcctcgtactatactcgtgaanaaacagcgcttaacatccagaaaacattgccc 1138
QY 974 acaacacgsgagagagctatctatgtgaaaagctcgaactggttcggagagcttacc 1033
Db 1139 tcggaacaaagagagagccctcctatgagcaaaaagctcctatagatcttaataatctaca 1198
QY 1034 ctattcagtgaaatagtgtaaatgaggtacagctcctaaagggtgagccatagctatc 1093
Db 1199 atatttcttgaaa-----tacagctgaaaaggcgctatgtctattc 1243
QY 1094 gaagatagtgsgaatgagttatctccgcgcatagtggtacatgtgctcttttagggat 1153
Db 1244 ccggaactcgtgagctcctcctcctacccgcaatloaagggtataltccctttaaacaagaac 1303
QY 1154 acagactcttactactcctcgtgagcaagatagaagtatcgactcagtaggaagagtgca 1213
Db 1304 c-----tcagcatcactagtgaggaaaccacacgcagcaatagatttcaacttcgaaaaagatgcc 1360
QY 1214 aagatgacacatttgctgttctgtgtgtgtagagccatctactctatgtatcccaact 1273
Db 1361 aagtttgcactctcctagggaaatacgcgaaggctataccctactactctatgtctccgattaca 1420
QY 1274 acggatcatcccaacaaagttacagatgctttaaagttatagactccggagattctc 1333
Db 1421 tctgagattattactctgtcatccgcgcgcgtacgtgtgctgcatacccaaggcagtt 1480
QY 1334 gca-----ctacaatatacagaagaaatcaatcctcacaagagaagaatatacagagacag 1390
Db 1481 gcagatggtgcgtatcagaggaactatgtctcttcagsgagaacccctcactgctacgcgaa 1540
QY 1391 gccgcgattctttaaatacttacttctgaagctactacagccttaactcttttagggagt 1450
Db 1541 gcagcaacccctcgaatgctacatcacaattaaaccaaaggctagaactgtgaagcggt 1600
QY 1451 actctactttaaacaatgtagtgaactcgtcagactcagactcagactcactcaacagcgagat 1510
Db 1601 actctcgtcttaagaacgggtgtactacttaaatggttcaactcgcgaagaatgaaag 1660
QY 1511 tctcgtcctgaaatagtagtgaactactctagaacctgctgata----- 1556
Db 1661 tccgtcgtcatcatagtagtgcagggaccacattagcaactacaatagagctataataact 1720
QY 1557 -----ctagacacataaacaatttgcgtatcaatacaatgcttctataagcgttgaaag 1609
Db 1721 gaagcgtcatcaacttaacaagaactgttaacaatcgtgactcttcttgatgagcattaa 1780
QY 1610 aaggcaaaaaatgaaacaaagctacgtacaaanaaatctgaacttaactctgaaacatcact 1669
Db 1781 gcggcgtcgttaatatgtgagagtaaccaatgtagcgtcctacatataccggaactttagaa 1840
QY 1670 ttatttgaccacagcggtcttttatgaaatcatagtttaagaatcctcagtcctac 1729
Db 1841 cttgtgaaaaaactctcaagattgctgtgcaaacacagcgatgtttaataaagattacag 1900
QY 1730 gac-----atcttagagctcaaaagcttctgga---acgttaacagaccccgcaatgact 1780
Db 1901 caagttcggatttttagaatacctcaaaagcgaacttcaataatcgtatacgaacttccagt 1960
```

```
QY 1781 ccagatccataataggtgtagaatactccattacgctcatcagsgaaacttgsgcccaatt 1840
Db 1961 ctggacacaacgagcatcagcatcactccctatggtatcaagaactcgtggagtttacc 2020
QY 1841 gtttggggacaggggctcttaacagctgaacacttaacactgaactaactgtctatatc 1900
Db 2021 atagacacgacaacccatcagctacagcaagaa-----aattgaaaaaaacggtatctc 2074
QY 1901 cctaaccggaggtatcggtctctttagtccttaagcttaagcttatgaaatcattatagat 1960
Db 2075 cccacacggagcgctctgctccctccttaacttaagccttatgggcaaacgtcatagat 2134
QY 1961 attagctctccatattattatgagactgcaaaaagaggggttcagaggagacggtgct 2020
Db 2135 ttacgagcgttaaggtacagcgtcagcagatgtatgtagaagatgccctcgggaagcaac-- 2192
QY 2021 tttgtgtgctgagatatactactcttcataaagatagtaacaaaacacagacgcggg 2080
Db 2193 -tgaatcacaggaattacaattcttccatgcaatcatalcaccggtgatgcagcagc 2251
QY 2081 ttccgcatattagtggtggttatgtacatagagaagaaacctcacaactgttccatataag 2140
Db 2252 taacgcataatggtgagagctacccacataactcaacacgcgcatcactcagaatgct 2311
QY 2141 attcctagtcgtcatattgtcagctcttggaaagatagtagactatttgaactaag 2200
Db 2312 gcgttaagcttaggttttgagacgctgttataaatactaaagatatactcgttaagtcac 2371
QY 2201 aatcaagtagacggtctacaggaagactctctatccagcacacaagaaacctatactct 2260
Db 2372 ggtcattcctaaggttatttgcctacagatatactatacatc----- 2413
QY 2261 ctctccttgaacctaagcgtctgtcgtgtgttattgttccacagagatccgtgtctc 2320
Db 2414 -----acaaagctcctggttgatatacgcgaattcttccaaaggacactcttc 2462
QY 2321 ttctcaggaacacttagctacacccatacgcgataacgaatctgtaaaaacaaatataca 2380
Db 2463 gag-----ttacctatagccgttagcaatggaagaagaaactctacatatcaaaa 2512
QY 2381 tatcctactgtttaaaggaagcgtgsgggaatgtagttccgctttagaattcgtgtagaa 2440
Db 2513 ttgcctaaagggcgtcgtcctctgtagtaacaattgtgttaagggaactcgaaggagac 2572
QY 2441 gctcgaattgcttagatgaaagtctcta---tttgagcagttacatgcccttaagaa 2497
Db 2573 ctcccatcaactctctctctcgcactcttaaacctcagaacgacatcctccttgtaaaa 2632
QY 2498 ttgcagtttgcctatgacacatcaggaaggttttaagaacagsgaagacgtcgtgaa 2557
Db 2633 gctgaagttgtttagcgcactatgtaggcatccaaagaaaataaccccgaggagagatt 2692
QY 2558 ttggaagtagccgctcgtgtgaatctcttaccctacacggagatccgattgtgaagaa 2617
Db 2693 ttggaacacgtlcatcactctcaacgtttgcaattcccgtaggcgtccgcttggtaaaaat 2752
QY 2618 tcaagctgcgaagatgcaacgtataaatctcttggtatctgtagtctgtgtcgt 2677
Db 2753 tctcataatcgacaggaatttttaacataataatcgttagcctatgtctatcgt 2812
QY 2678 agtaaccggagcttgaacacacacatcgtgaattagcgtgtgattcttggaaaacctcgtg 2737
Db 2813 caaatccctgatttcgatacgaacttaacttaatggaagtaagctagtgactctataggg 2872
QY 2738 aagaatttggaacagaacagctttagctccttgcgtgcagggagcaacatttcttataatca 2797
Db 2873 aataatcaacagaagtaacttctgtctagtaacagcaccagccatcacttaagtaaaagat 2932
QY 2798 aatttgaagcctttagccaatttcttcttgaaattgctgtgtgtatcgtcaattacaat 2857
Db 2933 gtctatagatcttcgggacactgtagatgtgatatctcgaagaacctcccgtaatatatact 2992
```







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 22:58:13 ; Search time 298.85 Seconds  
(without alignments)  
1518.234 Million cell updates/sec

Title: US-09-428-122-1

Perfect score: 3000  
1 cgccttacctagtagagst.....tgitttgctaacacttc 3000

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 262060 seqs, 75620727 residues

Total number of hits satisfying chosen parameters: 524120

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents: NA: \*  
1: /cgn2\_6/ptodata/2/1na/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/2/1na/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/2/1na/5C.COMB.seq: \*  
4: /cgn2\_6/ptodata/2/1na/5D.COMB.seq: \*  
5: /cgn2\_6/ptodata/2/1na/6.COMB.seq: \*  
6: /cgn2\_6/ptodata/2/1na/PCRN.COMB.seq: \*  
7: /cgn2\_6/ptodata/2/1na/Backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36.8	1.2	1269	3	US-08-758-621-11
2	35.8	1.2	1951	2	US-08-487-880A-112
3	35.8	1.2	1951	4	US-08-478-435-112
4	35.8	1.2	1951	4	US-08-337-483-112
5	35.8	1.2	1951	4	US-08-478-373-112
6	35.8	1.2	1951	5	US-08-474-671-112
7	35.8	1.2	1951	5	US-08-483-577A-112
8	35.6	1.2	8532	5	US-08-452-655B-1
9	35.6	1.2	8532	5	US-08-450-582-1
10	35.6	1.2	9606	1	US-07-741-940-1
11	35.6	1.2	9606	1	US-08-289-548A-1
12	35.6	1.2	9606	2	US-08-452-654-1
13	35.6	1.2	9606	2	US-08-370-235A-1
14	35.2	1.2	1120	1	US-08-254-493-3
15	35.2	1.2	1120	2	US-08-408-222B-3
16	35.2	1.2	2411	1	US-07-952-755-1
17	35.2	1.2	2411	2	US-08-443-679-1
18	34.2	1.1	8220	3	US-08-568-459A-11
19	34.2	1.1	8220	4	US-08-487-826B-11
20	34.2	1.1	19124	4	US-08-487-826B-13
21	33.4	1.1	2061	5	US-09-224-024-30
22	33.4	1.1	2061	6	PCT-US94-07902-30
23	33.4	1.1	4383	7	5175085-4
24	33.4	1.1	4383	7	5177307-1
25	33.2	1.1	3455	3	US-08-861-464-3
26	33.2	1.1	3455	4	US-08-396-001-3

27	33.2	1.1	7218	1	US-08-232-463-14	Sequence 14, Appl
28	32.8	1.1	3330	2	US-08-149-103-1	Sequence 1, Appl1
29	32.8	1.1	3330	2	US-08-451-883-1	Sequence 1, Appl1
30	32.8	1.1	9919	5	US-08-880-179-1	Sequence 1, Appl1
31	32.6	1.1	1604	2	US-08-694-579-1	Sequence 1, Appl1
32	32.6	1.1	1604	3	US-08-948-155-1	Sequence 1, Appl1
33	32.4	1.1	2694	5	US-08-975-703-5	Sequence 5, Appl1
34	32.2	1.1	2882	1	US-08-393-219-2	Sequence 2, Appl1
35	32.2	1.1	246240	3	US-08-724-394A-20	Sequence 20, Appl
36	32.2	1.1	246240	3	US-08-724-394A-21	Sequence 21, Appl
37	32.2	1.1	246240	3	US-08-724-394A-22	Sequence 22, Appl
38	32	1.1	347	1	US-08-149-695-3	Sequence 3, Appl1
39	32	1.1	347	1	US-08-377-228-3	Sequence 3, Appl1
40	32	1.1	377	1	US-08-149-695-4	Sequence 4, Appl1
41	32	1.1	377	1	US-08-377-228-4	Sequence 4, Appl1
42	32	1.1	464	1	US-08-149-695-5	Sequence 5, Appl1
43	32	1.1	464	1	US-08-377-228-5	Sequence 5, Appl1
44	32	1.1	595	1	US-08-149-695-6	Sequence 6, Appl1
45	32	1.1	595	1	US-08-377-228-6	Sequence 6, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-758-621-11  
: Sequence 11, Application US/08758621  
: Patent No. 5846821  
: GENERAL INFORMATION:  
: APPLICANT: Guerinet, Mary Lou, and Elide, David J.  
: TITLE OF INVENTION: Metal-Regulated Transporters and Uses Therefor  
: NUMBER OF SEQUENCES: 21  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: LAHIVE & COCKFIELD  
: STREET: 60 State Street  
: CITY: Boston  
: STATE: Massachusetts  
: COUNTRY: USA  
: ZIP: 02109-1875  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/758,621  
: FILING DATE:  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: 60/018,578  
: FILING DATE: 29-MAY-1996  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Silverl, Jean M.  
: REGISTRATION NUMBER: 39,030  
: REFERENCE/DOCKET NUMBER: DCI-099CP  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (617)227-7400  
: TELEFAX: (617)227-5941  
: INFORMATION FOR SEQ ID NO: 11:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 1269 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: MOLECULE TYPE: CDNA  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: 1..1267  
: US-08-758-621-11  
Query Match 1.2%; Score 36.8; DB 3; Length 1269;  
Best Local Similarity 49.0%; Pred. No. 0.46;  
Matches 98; Conservative 0; Mismatches 102; Indels 0; Gaps 0;



NAME/KEY: CDS  
LOCATION: 1..1890  
US-08-478-435-112

Query Match 1.2%; Score 35.8; DB 4; Length 1951;  
Best Local Similarity 52.3%; Pred. No. 1.1;  
Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

OY 30 ttcttgactgtttctccatgttgatctcttaaatatcaaatcaaga 89  
DB 1099 TTGTTGACTGCTGTTGCATCGGTTTGCATGCTTTTGTAGTACGAAAGTGTTA 1040  
OY 90 tatattacaatgaagtccttccccaagttgtatcttctacattgtatcc 149  
DB 1039 GCTTCCATCAATTAAGGTTCTTTGAGTAATTTTGTCTTCCTCCGTTTCTGCTGT 980  
OY 150 ctgtgtatgtatgtctaccgagacgttt 180  
DB 979 CTTTGCACTAATATACCCCAAAACTTTT 949

## RESULT 4

US-08-337-483-112/c  
Sequence 112 Application US/08337483  
Patent No. 3922362

GENERAL INFORMATION:  
APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Murdin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/337,483  
FILING DATE: 08-NOV-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 112:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1951 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1890  
US-08-337-483-112

Query Match

1.2%; Score 35.8; DB 4; Length 1951;

Best Local Similarity 52.3%; Pred. No. 1.1;  
Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

OY 30 ttcttgactgtttctccatgttgatctcttaaatatcaaatcaaga 89  
DB 1099 TTGTTGACTGCTGTTGCATCGGTTTGCATGCTTTTGTAGTACGAAAGTGTTA 1040  
OY 90 tatattacaatgaagtccttccccaagttgtatcttctacattgtatcc 149  
DB 1039 GCTTCCATCAATTAAGGTTCTTTGAGTAATTTTGTCTTCCTCCGTTTCTGCTGT 980  
OY 150 ctgtgtatgtatgtctaccgagacgttt 180  
DB 979 CTTTGCACTAATATACCCCAAAACTTTT 949

## RESULT 5

US-08-478-373-112/c  
Sequence 112 Application US/08478373  
Patent No. 5922841

GENERAL INFORMATION:  
APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Murdin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,373  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/337,483  
FILING DATE: 08-NOV-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 112:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1951 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1890

US-08-478-373-112

Query Match 1.2%; Score 35.8; DB 4; Length 1951;  
Best Local Similarity 52.3%; Pred. No. 1.1;  
Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 30 ttcttgactgttcttccatctgtatctcttaaaatcaaatcaaga 89  
DB 1099 TTGTTGACTGCTGTGTCATCGGTGTGTCATGTTTGTAGTAGAAGTACTTA 1040  
QY 90 tataattcaaatgaagctcttccccaagtgttatcttcaattgtcatttcc 149  
DB 1039 GCTTGCATCAATTAAGGTTCTTTGAGTAATTTTGTTCCTCCGTTCTTGCTGT 980  
QY 150 ctgtctatgatgtctaccgagacagttt 180  
DB 979 CTTTGCACTAATAATACCCCAAAACTTTT 949

RESULT 6

US-08-474-671-112/c  
Sequence 112, Application US/08474671  
Patent No. 6008326

GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Murdin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,671  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530

Prior Application Data:  
APPLICATION NUMBER: US 08/337,483  
FILING DATE: 08-NOV-1994  
CLASSIFICATION: 530

Prior Application Data:  
APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993

Prior Application Data:  
APPLICATION NUMBER: US 08/148,968  
FILING DATE: 08-NOV-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I

REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-465 MIS-V9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 112:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1951 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single

TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1890  
US-08-474-671-112

Query Match 1.2%; Score 35.8; DB 5; Length 1951;  
Best Local Similarity 52.3%; Pred. No. 1.1;  
Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 30 ttcttgactgttcttccatctgtatctcttaaaatcaaatcaaga 89  
DB 1099 TTGTTGACTGCTGTGTCATCGGTGTGTCATGTTTGTAGTAGAAGTACTTA 1040  
QY 90 tataattcaaatgaagctcttccccaagtgttatcttcaattgtcatttcc 149  
DB 1039 GCTTGCATCAATTAAGGTTCTTTGAGTAATTTTGTTCCTCCGTTCTTGCTGT 980  
QY 150 ctgtctatgatgtctaccgagacagttt 180  
DB 979 CTTTGCACTAATAATACCCCAAAACTTTT 949

RESULT 7

US-08-483-577A-112/c  
Sequence 112, Application US/08483577A  
Patent No. 6015688

GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Murdin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 160  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/483,577A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435

Prior Application Data:  
APPLICATION NUMBER: US 08/337,483  
FILING DATE: 08-NOV-1994

Prior Application Data:  
APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993

Prior Application Data:  
APPLICATION NUMBER: US 08/148,968  
FILING DATE: 08-NOV-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I

REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-511  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 112:  
SEQUENCE CHARACTERISTICS:

Tue Nov 21 11:10:26 2000

us-09-428-j22-1.rni

Page 5

```

:   LENGTH: 1951 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
:
:   FEATURE:
:     NAME/KEY: CDS
:     LOCATION: 1..1890
:
US-08-483-577A-112

```

Query Match	1.2%	Score 35.8	DB 5	Length 1951
Best Local Similarity	52.3%	Pred. No. 1.1		
Matches 79; Conservative	0	Mismatches 72	Indels 0	Gaps 0

Qy	30	ttcttgactggtttccctcttgatctcttaaatatcaatccaatcaagta	89
Db	1099	TTGATGACGGGTGTTCATCGGTGTCATGGTTTTTATAGAGAAAGNATA	1040
Qy	90	tatattacaagaagtcctcttcccaagttgattctcaatctgatatcc	149
Db	1039	GCTTGCATCAATTAAGGTTCTTTTAGNATTTTTTGTTCTTCCGTTCTTCTGTT	980
Qy	150	cttggtctatgatgtctaccgacagcttt	180
Db	979	CTTTGGCATTAAATACCCCAAAACTTTTTT	949

RESULT 8  
US-08-452-

US-08-452-655B-1/c  
; Sequence 1, Application US/08452655B  
; Patent No. 5782666

; GENERAL INFORMATION:

1 APPLICANT: ALBERTSEN, HANS  
 2 APPLICANT: ANAND, RAKESH  
 3 APPLICANT: CARLSON, MARY  
 4 APPLICANT: GRODEN, JOANNA  
 5 APPLICANT: HEDGE, PHILIP J.  
 6 APPLICANT: JOSLYN, GEOFF  
 7 APPLICANT: KINZLER, KENNETH  
 8 APPLICANT: MARKHAM, ALEXANDER F.  
 9 APPLICANT: NAKAMURA, YUSUKE  
 10 APPLICANT: THILIVERS, ANDREW  
 11 TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
 12 TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
 13 NUMBER OF SEQUENCES: 102  
 14 CORRESPONDENCE ADDRESS:

```

? TELEFAX: 202-580-9299
? INFORMATION FOR SEQ ID NO: 1
?:
?: SEQUENCE CHARACTERISTICS:
?:
?: LENGTH: 8532 base pairs
?: TYPE: nucleic acid
?: STRANDEDNESS: double
?: TOPOLOGY: linear
?: MOLECULE TYPE: cDNA
?: ORIGINAL SOURCE:
?: ORGANISM: Homo sapiens
?: IMMEDIATE SOURCE:
?: CLONE: DP2.5(APC)
US-08-452-655B-1

```

Query Match	1.2%	Score 35.6	DB 2	Length 8532
Best Local Similarity	57.0%	Pred. No. 3.1		
Matches 65; Conservative	0	Mismatches 49	Indels 0	Gaps 0

QY 1937 agcctatggaatcgcattatagatattagctccctccatctactctatgagacgccaac 1996  
 Db 760 AGCGGATTATGCTTTGTTGTGAGATGAGACTCTCTGCTCTCTGCTTGCGACTATAAA 701  
 QY 1997 gaagaggttcgaagagagacgcgtctcttttcggtgctgcgatcttctaactcttc 2050  
 Db 700 GCTTCTGATATACAGAGATATCTCTTTTTCGATTTGCGAATTTGCGATATCTTC 647

RESULT 9

US-08-450-582-1/c  
; Sequence 1, Application US/08450582  
; Patent No. 6114124

GENERAL INFORMATION:

1 APPLICANT: ALBERTSEN, HANS  
 2 APPLICANT: ANAND, RAKESH  
 3 APPLICANT: CARLSON, MARY  
 4 APPLICANT: GRODEN, JOANNA  
 5 APPLICANT: HEDGE, PHILIP J.  
 6 APPLICANT: JOSLYN, GEOFF  
 7 APPLICANT: KINZLER, KENNETH  
 8 APPLICANT: MARKHAM, ALEXANDER F.  
 9 APPLICANT: NAKAMURA, YUSUKE  
 10 APPLICANT: THILVERIS, ANDREW  
 11 TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
 12 TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
 13 NUMBER OF SEQUENCES: 102  
 14 CORRESPONDENCE ADDRESSES:



TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
CLONE: DP2.5(APC)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 34...8562  
US-08-289-548A-1

Query Match 1.2%; Score 35.6; DB 1; Length 9606;  
Best Local Similarity 57.0%; Pred. No. 3.3;  
Matches 65; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1937 agcttatggaatgcattatagatatagctctccattatctatggagactgcaaac 1996  
DB 793 ACCCGGTTTCATGCTTGTCTGAGATGACCTCTCTGCTTCTGCTGGAGACTGTAATA 734

QY 1997 gaagggttcaggagagacgctgtctttgtgtgctgattatctactcttc 2050  
DB 733 GCTGTCGATACGAGATGATGCTTTCGATTGCTGAATTCGGCTATTCCTTC 680

## RESULT 12

US-08-452-654-1/c  
Sequence 1, Application US/08452654  
Patent No. 5681454

## GENERAL INFORMATION:

APPLICANT: ALBERTSEN, HANS  
APPLICANT: ANAND, RAKESH  
APPLICANT: CARLSON, MARY  
APPLICANT: GRODEN, JOANNA  
APPLICANT: HEDGE, PHILIP J.  
APPLICANT: JOSLYN, GEORF  
APPLICANT: KINZLER, KENNETH  
APPLICANT: MARKHAM, ALEXANDER F.  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: THLIVERIS, ANDREW  
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner, Birch, McKie & Beckett  
STREET: 1001 G Street, NW  
CITY: Washington  
STATE: D.C.

COUNTRY: USA  
ZIP: 20001-4598

## COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/452,654  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/741,940  
FILING DATE: 08-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 1107.035574  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
SEQUENCE CHARACTERISTICS:

LENGTH: 9606 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
IMMEDIATE SOURCE:  
CLONE: DP2.5(APC)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 34...8562  
US-08-452-654-1

Query Match 1.2%; Score 35.6; DB 2; Length 9606;  
Best Local Similarity 57.0%; Pred. No. 3.3;  
Matches 65; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1937 agcttatggaatgcattatagatatagctctccattatctatggagactgcaaac 1996  
DB 793 ACCCGGTTTCATGCTTGTCTGAGATGACCTCTCTGCTTCTGCTGGAGACTGTAATA 734

QY 1997 gaagggttcaggagagacgctgtctttgtgtgctgattatctactcttc 2050  
DB 733 GCTGTCGATACGAGATGATGCTTTCGATTGCTGAATTCGGCTATTCCTTC 680

## RESULT 13

US-08-370-235A-1/c  
Sequence 1, Application US/08370235A  
Patent No. 5910418

## GENERAL INFORMATION:

APPLICANT: VOGELSTEIN, BERT  
APPLICANT: KINZLER, KENNETH W.  
APPLICANT: HILL, DAVID E.  
APPLICANT: JOHNSON, KAREN A.  
TITLE OF INVENTION: ANTIBODIES AND ASSAYS FOR DETERMINING  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BANNER & WITCOFF, LTD.  
STREET: 1001 G STREET, N.W.  
CITY: WASHINGTON  
STATE: DC

COUNTRY: US  
ZIP: 20001

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/370,235A  
FILING DATE: 01-JAN-1995  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: KAGAN, SARAH A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 01107.48688  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 508 9100  
TELEFAX: 202 508 9299  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9606 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:



```

; ORGANISM: Homo sapiens
; POSITION IN GENOME:
; MAP POSITION: 5q21
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 34..8562
; US-08-370-235A-1

Query Match          1.2%; Score 35.6; DB 4; Length 9606;
Best Local Similarity 57.0%; Pred. No. 3.3;
Matches 65; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1937 agcttaaggaatcattatagatagctctccattatcttaagagacgcaaac 1996
Db 793 AGCGGTTTCATGCTGTTCTGAGATGACCTCTGCTTGTGTTGGAGACTGTA 734
QY 1997 gaagggtgcagagagacgcgttttggtggtcgtgattatcttaattcttc 2050
Db 733 GCTGTCTATACGAGATGATGCTTTCGATTGCTGTAATCTGACTATTCTTC 680

RESULT 14
US-08-254-493-3
; Sequence 3, Application US/08254493
; Patent No. 5439886
; GENERAL INFORMATION:
; APPLICANT: IKEYAMA, SHUICHI
; APPLICANT: KOYAMA, MASARU
; APPLICANT: MIYAKE, MASAYUKI
; APPLICANT: SENOO, MASAHARU
; TITLE OF INVENTION: MONOCLONAL ANTIBODY, POLYPEPTIDES AND
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS & CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/254,493
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/865552
; FILING DATE: 09-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 079996-1991
; FILING DATE: 12-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 085396-1991
; FILING DATE: 17-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 022321-1992
; FILING DATE: 07-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: RESNICK, DAVID S.
; REGISTRATION NUMBER: 34235
; REFERENCE/DOCKET NUMBER: 41777
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; TELEX: 200281 STRE UR
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1120 base pairs
```

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to genomic RNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: HUMAN
; CELL TYPE: BREAST CARCINOMA
; CELL LINE: ZR-75-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 112..795
; NAME/KEY: mat _peptide
; LOCATION: 115..795
; US-08-254-493-3

Query Match          1.2%; Score 35.2; DB 1; Length 1120;
Best Local Similarity 57.1%; Pred. No. 1.2;
Matches 64; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 50 ttggtgatctcttaaaatcaatcaaaatcaagatatatttacaatgaagct 109
Db 951 TTATGTTTGTCTTTATGCTTCATTCATTAATGACATTGTGAGCGGGGTTT 1010
QY 110 tcttcccgaagtttgatttctacattgctatttcctcttgctatga 161
Db 1011 GCTTGTGTTGTTTATATTATTTTCAGTTGTTTGTTCCTTGTTATATA 1062

RESULT 15
US-08-408-222B-3
; Sequence 3, Application US/08408222B
; Patent No. 5776727
; GENERAL INFORMATION:
; APPLICANT: Ikeyama, Shuichi
; APPLICANT: Koyama, Masaru
; APPLICANT: Miyake, Masayuki
; APPLICANT: Senoo, Masaharu
; TITLE OF INVENTION: MONOCLONAL ANTIBODY, POLYPEPTIDES AND
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, Bronstein, Roberts & Cushman
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Pasteo Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/408,222B
; FILING DATE: 22-MAR-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/254,493
; FILING DATE: 06-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP-079996-1991
; FILING DATE: 12-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP-085396-1991
; FILING DATE: 14-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP-022321-1992
; FILING DATE: 07-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S.
```

REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 41777-DIV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1120 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to genomic RNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM: human  
CELL TYPE: breast carcinoma  
CELL LINE: ZR-75-1  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 112...795  
OTHER INFORMATION: E Mat peptide  
US-08-408-222B-3

Query Match 1.2%; Score 35.2; DB 2; Length 1120;  
Best Local Similarity 57.1%; Pred. No. 1.2;  
Matches 64; Conservative 0; Mismatches 48; Indels 0; Gaps 0;  
QY 50 ttgggtatctcttaataataatcaaatcaaatcaagatatatttacaatgaagtct 109  
DB 951 ttatgtttgtctttaatgcttcattcaatataatgacattgtgagcgcgcggttt 1010  
QY 110 tcttccccaaattgatttctcaatttgctatttcccttgatga 161  
DB 1011 ggttgcttggttatatatttttcagtgctgttttgccttgatatata 1062

Search completed: November 20, 2000, 09:20:50  
Job time: 37357 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 19, 2000, 20:42:58 ; Search time 1452.66 Seconds  
(without alignments)  
12772.107 Million cell updates/sec

Title: US-09-428-122-1

Perfect score: 3000

Sequence: 1 cgctctaccctagtagaggt.....tgcttgcctaaacacttic j000

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7189864 seqs, 3091403243 residues

Total number of hits satisfying chosen parameters: 14379728

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*\n2: gb\_est2:\*\n3: gb\_est3:\*\n4: gb\_est4:\*\n5: gb\_est5:\*\n6: gb\_est6:\*\n7: gb\_est7:\*\n8: gb\_est8:\*\n9: gb\_est9:\*\n10: gb\_est10:\*\n11: gb\_est11:\*\n12: gb\_est12:\*\n13: gb\_est13:\*\n14: gb\_est14:\*\n15: gb\_est15:\*\n16: gb\_est16:\*\n17: gb\_est17:\*\n18: gb\_est18:\*\n19: gb\_est19:\*\n20: gb\_est20:\*\n21: gb\_est21:\*\n22: gb\_est22:\*\n23: gb\_est23:\*\n24: gb\_est24:\*\n25: gb\_est25:\*\n26: gb\_est26:\*\n27: gb\_est27:\*\n28: gb\_est28:\*\n29: gb\_est29:\*\n30: gb\_est30:\*\n31: gb\_est31:\*\n32: gb\_est32:\*\n33: gb\_est33:\*\n34: gb\_est34:\*\n35: gb\_est35:\*\n36: gb\_est36:\*\n37: gb\_est37:\*\n38: gb\_est38:\*\n39: gb\_est39:\*\n40: gb\_est40:\*\n41: em\_estda:\*\n42: em\_estfun:\*\n43: em\_esthum1:\*

44: em\_esthum2:\*\n45: em\_esthum3:\*\n46: em\_esthum4:\*\n47: em\_esthum5:\*\n48: em\_esthum6:\*\n49: em\_esthum7:\*\n50: em\_esthum8:\*\n51: em\_esthum9:\*\n52: em\_esthum10:\*\n53: em\_esthum11:\*\n54: em\_esthum12:\*\n55: em\_esthum13:\*\n56: em\_esthum14:\*\n57: em\_esthum15:\*\n58: em\_esthum16:\*\n59: em\_esthum17:\*\n60: em\_esthum18:\*\n61: em\_esthum19:\*\n62: em\_esthum20:\*\n63: em\_estin1:\*\n64: em\_estin2:\*\n65: em\_estin3:\*\n66: em\_estin4:\*\n67: em\_estom:\*\n68: em\_estov1:\*\n69: em\_estov2:\*\n70: em\_estp1:\*\n71: em\_estp12:\*\n72: em\_estp13:\*\n73: em\_estp14:\*\n74: em\_estp15:\*\n75: em\_estro1:\*\n76: em\_estro2:\*\n77: em\_estro3:\*\n78: em\_estro4:\*\n79: em\_estro5:\*\n80: em\_estro6:\*\n81: em\_estro7:\*\n82: em\_estro8:\*\n83: em\_estro9:\*\n84: em\_estro10:\*\n85: em\_estro11:\*\n86: em\_estro12:\*\n87: em\_estro13:\*\n88: gb\_gsa1:\*\n89: gb\_gsa2:\*\n90: gb\_gsa3:\*\n91: gb\_gsa4:\*\n92: em\_gsa1:\*\n93: em\_gsa2:\*\n94: em\_gsa3:\*\n95: em\_gsa4:\*\n96: gb\_gsa5:\*\n97: gb\_gsa6:\*\n98: gb\_gsa7:\*\n99: gb\_gsa8:\*\n100: gb\_gsa9:\*\n101: em\_gsa5:\*\n102: em\_gsa6:\*\n103: em\_gsa7:\*\n104: em\_gsa8:\*\n105: em\_gsa9:\*\n106: em\_gsa10:\*\n107: em\_gsa11:\*\n108: gb\_gsa10:\*\n109: gb\_gsa11:\*\n110: em\_gsa12:\*\n111: gb\_gsa12:\*\n112: gb\_gsa13:\*\n113: gb\_gsa14:\*\n114: gb\_gsa15:\*\n115: gb\_gsa16:\*\n116: gb\_gsa17:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	50.6	1.7	923	121	CNS01509	AL105699 Drosophila
C 2	49.4	1.6	334	15	AU076383	AU076383 Drosophila
C 3	44.4	1.5	1025	121	CNS01492	AL104216 Drosophila
C 4	43.8	1.5	596	124	FR0019514	AL012404 F. rubripes
C 5	43.2	1.4	1101	121	CNS00871	AL052689 Drosophila
C 6	42.8	1.4	783	121	CNS00A15	AL055883 Drosophila
C 7	42.8	1.4	1201	121	CNS0167M	AL105396 Drosophila
C 8	42.4	1.4	240	15	AU072820	AU072820 Drosophila
C 9	42.4	1.4	1101	121	CNS0039G	AL065921 Drosophila
C 10	42.2	1.4	900	108	AU0744776	AU0744776 HS_5506_2
C 11	41.6	1.4	553	19	AM019090	AM019090 f439g04.x
C 12	41.6	1.4	1101	121	CNS00K8B	AL077903 Drosophila
C 13	41.4	1.4	1101	121	CNS017NE	AL108260 Drosophila
C 14	41.2	1.4	315	27	BI126193	BI126193 Drosophila
C 15	41	1.4	529	34	BE225902	BE225902 fa21c12.x
C 16	41	1.4	735	133	CNS04NSM	AL1299119 Tetradon
C 17	41	1.4	1309	35	BE420736	BE420736 HMM002.AO
C 18	40.8	1.4	494	89	AO124208	AQ124208 HS_3122_B
C 19	40.8	1.4	564	23	AM643885	AM643885 cm35a01.w
C 20	40.8	1.4	566	111	AQ829160	AQ829160 HS_3112_E
C 21	40.8	1.4	639	121	CNS0170D	AL108367 Drosophila
C 22	40.8	1.4	949	121	CNS00CJN	AL059162 Drosophila
C 23	40.8	1.4	919	121	CNS00JDN	AL076146 Drosophila
C 24	40.8	1.4	1101	121	CNS00E16	AL067654 Drosophila
C 25	40.6	1.4	242	15	AU037868	AU037868 Drosophila
C 26	40.6	1.4	959	121	CNS00655	AL062806 Drosophila
C 27	40.6	1.4	975	121	CNS010XM	AL099595 Drosophila
C 28	40.6	1.4	1101	121	CNS002DG	AL0662437 Drosophila
C 29	40.2	1.3	652	118	BI5488	BI5488 345E3_TV_C1
C 30	40.2	1.3	964	122	CNS02GJO	AL156413 Tetradon
C 31	40.2	1.3	1201	121	CNS0163T	AL106255 Drosophila
C 32	40	1.3	524	122	CNS0109U	AL167544 Tetradon
C 33	40	1.3	577	14	AI957567	AI957567 fc94h04.x
C 34	40	1.3	942	121	CNS018B5	AL109338 Drosophila
C 35	40	1.3	1101	121	CNS016L1	AL110633 Drosophila
C 36	39.8	1.3	905	121	CNS00KHX	AL1077798 Drosophila
C 37	39.8	1.3	940	122	CNS02GSH	AL196730 Tetradon
C 38	39.8	1.3	988	121	CNS00D7R	AL066743 Drosophila
C 39	39.8	1.3	1101	121	CNS00077	AL075253 Drosophila
C 40	39.8	1.3	1101	121	CNS000Z2H	AL0797139 Drosophila
C 41	39.8	1.3	1101	121	CNS0100P	AL0589371 Drosophila
C 42	39.8	1.3	1176	123	CNS03J83	AL265908 Tetradon
C 43	39.6	1.3	1028	123	CNS04RPD	AL304186 Tetradon
C 44	39.6	1.3	1034	121	CNS0154I	AL104988 Drosophila
C 45	39.6	1.3	1101	121	CNS001GJ	AL060877 Drosophila

## ALIGNMENTS

LOCUS	NS01509	923 bp	DNA	SSS	26-JUL-1999
DEFINITION	Drosophila melanogaster genome survey sequence Sp6 end of BAC BACH14B20 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL105699.1 GI:5617836				
VERSION	SSS.				
KEYWORDS	fruit fly.				
SOURCE	Drosophila melanogaster				
ORGANISM	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 923)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - FRANCE (E-mail : secrete@genoscope.cns.fr)				
COMMENT	<p>Web : www.genoscope.cns.fr</p> <p>Determination of this BAC-end sequence was carried out as part of the collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.</p>				
FEATURES	<p>source</p> <p>location=Qualifiers</p> <p>1..923</p> <p>/organism="Drosophila melanogaster"</p> <p>/plasmid="pBeloBAC11"</p> <p>/db_xref="taxon:7227"</p> <p>/clone_11b="DrosBAC"</p> <p>/clone="BACN14B20"</p> <p>/note="end : SP6"</p>				
BASE COUNT	245 a 142 c 100 g 315 t 121 others				
ORIGIN					
Query Match	1.7%; Score 50.6; DB 121; Length 923;				
Best Local Similarity	34.5%; Pred. No. 0.0081;				
Matches 129; Conservative	66; Mismatches 178; Indels 1; Gaps 1;				
QY	29 ttctctgactgtctctcctactgtgctatctctctataatataatcaatcaagt 88				
Db	909 TTTTATTTTTTTTAAATWYATTTTATATTTTAAATATTTATTAATTTT 850				
QY	89 atattttacatgaagtctctctccccaag-tttgtattttctacattgtcat 147				
Db	849 TTWTWTTTMMWAAADTTTAAATKHHWAAGBTKYTTTBTTTTATAGAGGATT 790				
QY	148 ccccttgctatgattgctacgacagattttgattatcaagtgcgagttcagtgaggaa 207				
Db	789 YCMRKADNSNNNSNMAMMMHMASSSSVSCAAKAKGAYVCGMGMRKRAMHINVGA 730				
QY	208 taataatgtaattttcaagtcgtgagatgcacagaagaatgctggaactacactcatt 267				
Db	729 AAAAAMAAAANAAYVANDKNCSSGGNAGGGGSSBSRGAATMGCGACAKTBTYAT 670				
QY	268 taagggaatgtcactcagaataatctccggacacagcagcagcaatcacaaaaagctg 327				
Db	669 MGGAMACAAAMAAAAMAAAAMAAAAGCAACWKAATAAAGTCTWHAATAACAAAAGAA 610				
QY	328 tttaacacacactaaggcgattgacttccacagtaaggaactctcatgttcca 387				
Db	609 ATMMABNBKBTBKBAATAATTGTCATTGAATCATTTWTTGAAAAAGAAAGGGGTTCTA 550				
QY	388 aacggtgtagtcag 401				
Db	549 AGAGTGGGGGAAG 536				

LOCUS	334 bp	mRNA	EST	27-JUL-1999
DEFINITION	AU076383	Dictyostelium discoideum SS (H.Urushihara)	Dictyostelium	
ACCESSION	AU076383	dictoideum cdna clone SSA686, mRNA sequence.		
VERSION	AU076383.1	GI:5607281		
KEYWORDS	EST.			
SOURCE	Dictyostelium discoideum.			
ORGANISM	Dictyostelium discoideum.			
AUTHORS	Eukaryota: Dictyostelidida, Dictyostelium.			
TITLE	1 (bases 1 to 334)			
COMMENT	Urushihara,H. Developmental cDNA in Dictyostelium discoideum (1999) Unpublished (1999) Contact: Hideko Urushihara Institute of Biological Sciences University of Tsukuba 3-3-10 Ten-noda, Tsukuba, Ibaraki 305, Japan Email: dd02huesakura.cc.tsukuba.ac.jp PROJECT = Dictyostelium discoideum cdna project in Japan. Location/Qualifiers			
FEATURES	1..334			
source	/organism="Dictyostelium discoideum"			
	/strain="AX4"			
	/db_xref="taxon:44689"			
	/clone="SSA686"			
	/clone_id="Dictyostelium discoideum SS (H.Urushihara)"			
	/dev_stage="slug"			
BASE COUNT	173 a	30 c	38 g	93 t
ORIGIN				
Query Match	1.6%	Score 49.4;	DB 15;	Length 334;
Best Local Similarity	48.1%;	Pred. No. 0.012;		
Matches 140;	Conservative 0;	Mismatches 151;	Indels 0;	Gaps
QY 29	ttcttgactggttcctccatgigtatcctctaataatcaatcaatcaagt 88			
DB 325	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAATTTTATT 266			
QY 89	ataattacaagaagtcctctccccaagttgattctctacattgctattc 148			
DB 265	TTTTATTTTATTTATTTTATTTTATTTTCTTAAATTTTCTTGAGTATTC 206			
QY 149	ccttgctcagatgctaccggaacagtttgattcaagtgcgaattcgtgggaat 208			
DB 205	CTATTTTGAATATTCGTCTACACCTTATTAATTCGAAATGTTTTCACAAAGATCT 146			
QY 209	aaaatgtaatttccttcagtcgtagagtcgaagaagtcgtgaactactcattt 268			
DB 145	AAATGTGGAACTTACCATTTCTAATAAGATCAGACACAATGTATAAATTAACCAATG 86			
QY 269	aagggaatgtcactctagaanaatatctctcgtgaacagcgacagcaatcaca 319			
DB 85	TATTCATTAAGCAACATTAACATTAATCAAGCCAAAGTGAAGATTAACA 35			
RESULT 3				
LOCUS	CNS014J2	1025 bp	DNA	SSS
DEFINITION	Dictyostelium melanogaster genome survey sequence T7 end of BAC BACN1111 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.			
ACCESSION	AL104216			
VERSION	AL104216.1	GI:5615827		
KEYWORDS	GSS.			
SOURCE	fruit fly.			
ORGANISM	Drosophila melanogaster			
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
REFERENCE	1 (bases 1 to 1025)			
AUTHORS	Genoscope.			
TITLE	Direct Submission			

[illegible]

Query Match		1.4%	Score 43.8;	DB 124;	Length 596;	
Best Local Similarity		45.5%;	Pred. No. 0.51;			
Matches 135; Conservative		0;	Mismatches 162;	Indels	0;	Gaps 0;

  

QY	540	gtacgtttagtttgacaataatgcatcagtttccttcacgaacaaactttcaaggata	599
Dd	244	GTAGTCTCAAGAACAACAGCACAAGGAANAGCANNTGGCTATGACTGATCTCGGATC	303
QY	600	atggcggtgcataccaccgcaaaaacttcattcaaacaggagatacaatgtaagctctgt	659
Dd	304	CCTGTAAGTCTGTTACTGCTGCTACTACTGCTACTGCTACTGCTACTGCTACTGCTACTA	363
QY	660	tttctgaaataactcctctaagaaggaggccattcaagattccgtgccttaccca	719
Dd	364	CCTCTACTGCTGTTACTGCTACTACTGCTGCTGCTGCTGCTGCTGCTACTGCTACTACG	423
QY	720	ttactgaaaccaagggaagtcctcttcttcacataacttccttgagtctgagcty	779
Dd	424	CTGCTGCTACTAAGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	483
QY	780	caattttcacagaagcctcggtgataatttcaataatgctaagttccattatgy	836
Dd	484	CTGCTGCTGCNCGTGCTACTACNACNGCTGCTACTGCTACTGCTGCTGCTGCTACTACG	540

  

RESULT	5
CNS008VI	
LOCUS	1101 bp DNA GSS 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence TEJ3 end of BAC #
ACCSSION	BAC1814 of RPCL-98 library from Drosophila melanogaster (fruit
VERSION	AL052689
KEYWORDS	AL052689.1 GI:4934238
SOURCE	GSS:
ORGANISM	fruit fly.
REFERENCE	Drosophila melanogaster
AUTHORS	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
JOURNAL	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
COMMENT	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

  

Genoscope.  
Direct Submission  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequenage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see http://www.fruitfly.org The BDGP Drosophila  
melanogaster BAC library was prepared by Kazutoyo Osogawa and  
Aaron Mammeter in Pieter de Jong's laboratory in the Department of  
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
NY. The library is named RPCL-98 and was constructed by partial  
EcoRI digestion of Drosophila DNA provided by the BDGP from the  
isogenic strain Y2: cn bw sp, the same strain used for the BDGP's  
P1 and EST libraries. A more detailed description of the library  
and how to order individual BAC clones, the entire library, or  
filters for hybridization from the BACPAC Resource Center can be  
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.

  

FEATURES	
source	1..1101
	/organism="Drosophila melanogaster"
	/db_xref="taxon:7227
	/clone_1lb="RPCL-98"
	/clone="BACR1814"
	/note="end : TEJ3"

  

BASE COUNT	289 a	112 c	113 g	316 t	271 others
ORIGIN					

  

Query Match 1.4%; Score 43.2; DB 121; Length 1101;

	Best Local Similarity	32.1%; Pred. No. 0.9;	Mismatches	125;	Indels	7;	Gaps	1;
	Matches	97;	Conservative	73;				
QY	36	actgtttccctatcattgtagtatacctctaanaatatlaaatccaagaatataatt	95					
		: : : : :   : :   :   :   :   :   :   :						
Db	684	ATTTSSHMSWSTTTAASMSMTTTTWTVATAAGMWAMWTWTAASAATTAAMAATSMM	743					
QY	96	ttaacaatgaagcttcctctcccacaaagttagtaatttcaacattgatctttacccttg	155					
		: : : : :   : :   : :   : :   : :   : :   :						
Db	744	TTAAAAATTSDDCTSTIAWSAVVMAWMTVVAALPAHSAGANVSKKKAWMAPACHTTISC	803					
QY	156	catacatygcacccagagacagitttggaatlccaagtgcagtltcgatyygaataaata	215					
		: :   :   :   :   :   :   :   :   :   :   :						
Db	804	HTTTTTTTAAATATAAAATGSGTSTAATATTTCGCYSTTAAWTVMGGRAMAALMWT	863					
QY	216	gtaattttcaattc-----gtgagagtcgaagaatgctcggaacctaacctattt	268					
		:   :   :   :   :   : : : : : : : : :   :						
Db	864	VKSTSMGTGSGITSSGKTWGKGKRWMTTSTAISTTKRTBSVVTVAAMSSYSTTTMT	923					
QY	269	aaggaggaatgctactctagaanaatctctcgyaacagycacagccatcaccaaaagcct	328					
		: :   :   :   :   :   :   :   :   :   :   :						
Db	924	TTSNATRTRSTFKKAKATGTTAAADTTCOCTAAMAGSGGAIAAAAAAAAAAAAAATK	983					
QY	329	tt 330						
Db	984	TT 985						

[illegible]

Query Match	1.4%	Score 42.8	DB 121	Length 783
Best Local Similarity	42.8%	Pred. No. 1	Mismatches 71	Indels 0
Matches 71	Conservative 24	Pred. Matches 71	Indels 0	Gaps 0
Oy	20	ttgagtgaaattcttgactggtcttcctcactgggtgatctcttaaatattcaatcaa	79	
Db	316	TTATTTTAATATTTTGGTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT	257	
Oy	80	aatcaagatataatttacaatgaagtccttcctcccaagttgatttctccattt	139	
Db	256	ATMMTTTTTTTTTTTTTATATWAGTCTTTTWTWTWTWTWTWTWTWTWTWTWT	197	
Oy	140	gctatttccttcctgcatgattgctacaggagacagtttgatt	185	
Db	196	TTCTTTTTTTTTTTTTTCCCKMTTTTTTCYTMMWTTKCAAT	151	
RESULT	7			
LOCUS	CNS0167M			
DEFINITION	CNS0167M 1201 bp DNA GSS 26-JUL-1999			
DESCRIPTION	Drosophila melanogaster genome survey sequence T7 end of BAC BACN15M24 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.			
ACCESSION	AL106396			
VERSION	ALI06396.1 GI:5621701			
KEYWORDS	GSS.			
SOURCE	fruit fly.			
ORGANISM	Drosophila melanogaster			
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
AUTHORS	1 (bases 1 to 1201)			
TITLE	Genoscope.			
JOURNAL	Direct Submission			
COMMENT	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)			
FEATURES	Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billard at CDPH (Centre d'Etude du Polymorphisme Humain) with fundings provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.			
SOURCE	Location/Qualifiers			
	1..1201			
	/organism="Drosophila melanogaster"			
	/plasmid="pBelobAC11"			
	/db_xref="taxon:7227"			
	/clone_lib="DrosBAC"			
	/clone="BACN15M24"			
	/note="end : T7"			
BASE COUNT	323 a 87 c 79 g 551 t 161 others			
ORIGIN				
Query Match	1.4%	Score 42.8	DB 121	Length 1201
Best Local Similarity	38.7%	Pred. No. 1.2	Mismatches 169	Indels 0
Matches 125	Conservative 29	Pred. Matches 169	Indels 0	Gaps 0
Oy	20	ttgagtgaaattcttgactggtcttcctcactgggtgatctcttaaatattcaatcaa	79	
Db	725	TTGTTTATGNAATTAAWTTATATTAAWTTTTTTTMMWTWTWTWTWTWTWTWT	794	
Oy	80	aatcaagatataatttacaatgaagtccttcctcccaagttgatttctccattt	139	
Db	795	AARATTTAAATTTTTTTTTTMAAAATWTTATATWAAATTTTTTMMTTTTTTTT	854	
Oy	140	gctatttccttcctgcatgattgctacaggagacagtttgattcagtcgagtttc	199	
Db	855	TTTTTTTWTWTWTATATATTTTAAWTTTTTTTTTAAWTTTTTTTTTAAWRAAA	914	

Oy	200	gatsggaaataaaaatgcgaatttttcagctcgtgtagagtcacgaagaatgcgaactacc	259
Db	915	TTTATTAATATAAAAAAAWTTTATATTWTTTTTTTTTTTTTTAAAAAAMWMTMAAAAATTTT	974
Oy	260	tacctattaaaggsgaaatgcacccctagaanaatttccctgsgaaacggacgcacatcca	319
Db	975	AAAAATTTTMTATTAATTATTATTATTTTAAAAATTTTTTTTTTTTWTWTATATAAAAAAAA	1034
Oy	320	aaaaatcgltttaacaacactaa	342
Db	1035	AAAAAAAAATTTATATNMAAAWMAA	1057
RESULT	8		
AU072820/c			
LOCUS	AU072820	240 bp	mRNA EST
DEFINITION	AU072820 Dictyostelium discoidem SS (H.Urushihara)		24-JUN-1999
ACCESSION	AU072820		
VERSION	AU072820.1	GI:5179241	
KEYWORDS	EST.		
SOURCE	Dictyostelium discoidem.		
ORGANISM	Dictyostelium discoidem.		
REFERENCE	Eukaryota: Dictyostellida; Dictyostelium.		
AUTHORS	Urushihara,H.		
TITLE	Developmental cDNA in Dictyostelium discoidem (1999)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Hideko Urushihara Institute of Biological Sciences University of Tsukuba 3-3-10 Ten-noda, Tsukuba, Ibaraki 305, Japan Email: d402nu@esakura.cc.tsukuba.ac.jp PROJECT = dictyostelium discoidem cDNA project in Japan. Location/Qualifiers 1..240 /organism="Dictyostelium discoidem" /strain="AX4" /db_xref="taxon:44689" /clone="SSF174" /clone_1fb="Dictyostelium discoidem SS (H.Urushihara)" /dev_stage="slug"		
FEATURES	source		
BASE COUNT	138 a	23 c	5 g 74 t
ORIGIN			
Query Match	1.4%; Score 42.4; DB 15; Length 240;		
Best Local Similarity	51.0%; Pred. No. 0.91;		
Matches 100:	Conservative 0; Mismatches 96; Indels 0; Gaps 0;		
Oy	28	attctctgaacttgcttcctccattcgggtgatcccttaaatatatnaattcaaatcaaag	87
Db	240	ATAGTGTCATGCTTTTATTATTATTATTATTATTTTAAAAAATTTAAAGCAGTTTATT	181
Oy	88	tatatatttcaatgaagcttccttcccacgaagttagtatcttcatattgctattt	147
Db	180	TATTTATTATAAAAAATTTTATTATTATTATTATTATTATTATTATTATTATTAAAAATTT	121
Oy	148	coccttgtctatgatctgacacagacagcttggatccaagtgcgaagtttcgatgsgaa	207
Db	120	TTTTTTTTTTTTTTTTTGCATGAAGGATATTTTAAAGGTGTGATACACATTTAG	61
Oy	208	taaaaatgctaattt	223
Db	60	TTACAATATGTGTTTT	45
RESULT	9		
CNS00396			
LOCUS	CNS00396	1101 bp	DNA GSS
DEFINITION	Drosophila melanogaster genome survey sequence TERT3 end of BAC #		03-JUN-1999
BACROBK10 of Rpci-98 library from Drosophila melanogaster (fruit			



Accession	Length	Strand	Gene	Feature	Version
AW019090	553 bp	+	zebrafish	EST	13-SEP-1999
ids590a.x1		+	zebrafish	EST	
sequence.		+	zebrafish	EST	
AW019090		+	zebrafish	EST	
AW019090.1	GI:5872620	+	zebrafish	EST	
Danio rerio		+	zebrafish	EST	



Oy	276	atgtcaccttaagaataatt	295 
Dd	639	CTKCCATTATGTCTTACT	658
RESULT 13			
CNS017NE	1101 bp	DNA	GSS      26-JUL-1999
LOCUS			
DEFINITION	Drosophila melanogaster genome survey sequence SpE end of BAC BAEN37L20 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
ACCESSION	AL108260		
VERSION	GI:5628564		
KEYWORDS			
SOURCE	fruit fly;		
ORGANISM	Drosophila melanogaster		
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
REFERENCE	Genoscope. Direct Submission Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)		
AUTHORS	- Web : www.genoscope.cns.fr) dermatation of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.		
JOURNAL			
COMMENT			
FEATURES			
Source	Location/Qualifiers		
	. . . 1101 /organism="Drosophila melanogaster" /plasmid="pBelobAC11"		
	/db_xref="taxon:7227"		
	/clone_lib="DrosBAC"		
	/clone="BAEN37L20"		
	/note="end : SP6"		
BASE COUNT	282 a    131 c    108 g    374 t    206 others		
ORIGIN			
Query Match	1.4% ; Score 41.4 ; DB 121 ; Length 1101 ; Best Local Similarity 46.8% ; Pred.No.2.8 ; Matches 81 ; Conservative 12 ; Mismatches 80 ; Indels 0 ; Gaps 0 ;		
OY	31	tcttacttggttcaccatgagtgcacccaaatatcaaatccaagcgaat	90                                   :
Dd	107	TTTTTTT TTTTTTT TTTTTTT TTATAAMWMTAATTTTTTTTTTAATAATMW	166                                   :
OY	91	ataattacaatgaagctcttctccccaaagttgatatttctaatttgtcatlccc	150                                     :
Dd	167	AATTTTTAAWTAGTGTTTTTATACCDCACTTAGTCCTTTTWTAATTCAATCWMAAT	226                                   :
OY	151	tttgctactgatctcgacggacaggitttgatccoagtgtaggttcgaag	203                                     :
Dd	227	TG TGTTAAAATGGAAATTTAATTTTNGCACAATTTTNCAAAMWWMGATG	279                                     :
RESULT 14			
BH126193	315 bp	mRNA	EST      28-JUN-2000
LOCUS			
DEFINITION	BH126193 RIKEN full-length enriched, 16 days neonate cerebellum musculus CDNA clone 9630012E11 3', mRNA sequence.		
ACCESSION	BH126193		
VERSION	BH126193.1 GI:8780528		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		

REFERENCE	Mammalia: Eutheria: Rodentia: Sotungnath: Muridae: Murinae; Mus. Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Hirozone, T., Horii, F., Ishii, Y., Ishikawa, D., Ishikawa, T., Itoh, M., Iizawa, N., Kadoya, K., Kawaga, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Salto, H., Sakai, C., Sato, K., Shibata, K., Shibata Y., Shimemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomihada, N., Toyota, T., Tsunoda, Y., Watanishi, A., Watanabe, S., Yamamura, T., Yamane, K., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino , M., Muramatsu, M. and Hayashizaki, Y.
TITLE	Riken Mouse ESis (Konno, H., et al.)
JOURNAL	Unpublished (2000)
COMMENT	Contact: Yoshihide Hayashizaki Genome Exploration Research Group, Life Science Tsukuba Center, Genome Science Laboratory The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel.: +81-298-36-9013 Fax: +81-298-36-9098 Email: genome-res@tc.riken.go.jp/ URL: http://genome.rtc.riken.go.jp/ Carninci, P., Miyajima, Y., Westover, A., Itoh, M., Nagaoaka, S., Sasaki , N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Thermoisabilization and thermostabilization of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998) Itoh, M., Kitsuai, T., Akizawa, Y., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki , Y. and Hayashizaki, Y. Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999) Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999) Please visit our web site ( <a href="http://genome.rtc.riken.go.jp">http://genome.rtc.riken.go.jp</a> ) for further details.
FEATURES	Location/Qualifiers
SOURCE	1..315 /organism="Mus musculus" /db_xref="taxon:10090" /clone_id="9630012E11" /clone_lib="RIKEN full-length enriched, 16 days neonate cerebellum" /tissue_type="cerebellum" /dev_stage="16 days neonate" /lab_host="DH10B" /note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCACAGCTTTTTCCTTTTTTTTNN 3'] cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 370.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCCGACTTAATTAATTAATCACCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT	94 a 53 c 33 g 135 t
ORIGIN	



Tue Nov 21 11:10:32 2000

us-09-428-122-1.rst

---

Page 10